

# GRID MOVIES

MATTHEW GRAHAM

**ABSTRACT.** We present a grid diagram analogue of Carter, Rieger and Saito's smooth movie theorem. Specifically, we give definitions for grid movies, grid movie isotopies and present a definition of grid planar isotopy as a particular subset of the grid diagram moves: stabilization, destabilization and commutation. We show that grid planar isotopy classes are in 1-1 correspondence with smooth planar isotopy classes by using a new planar grid algorithm that takes a smooth knot diagram to a grid diagram. We then present generalizations of both the smooth and grid movie theorems that apply to surfaces with boundary.

## 1. INTRODUCTION

The motivation for this topological work came from knot Floer homology. We briefly sketch the chain of ideas leading up to the questions addressed in this paper presently.

In 2000 Ozsváth and Szabó [14] introduced Heegaard Floer homology (HF). HF is a very powerful theory that provides a unified approach to studying 3- and 4-manifolds and the knots and links within them. It comes in several flavors: the hat version  $\widehat{HF}$  is the simplest; the minus and plus versions  $HF^-$ ,  $HF^+$  contain more geometric information; and there are other versions. Knot Floer homology (HFK), the version used to study knots and links was defined by Ozsváth and Szabó [13] and independently Rasmussen [15]. It uses Heegaard diagrams and holomorphic disks and is an invariant of oriented links in the 3-sphere, which, among other things, detects the genus of a link  $L$ , determines if  $L$  is fibered, and categorifies the Alexander polynomial. That is, the graded Euler characteristic of  $HFK(L)$  is the Alexander polynomial of  $L$ . A combinatorial method of calculating knot Floer homology was presented in [9, 11]. In [10], a completely combinatorial construction of HFK, independent of holomorphic techniques, was given. These methods use a grid diagram representation of an oriented knot as input. A *grid diagram* of *grid index*  $n$  is an  $n \times n$  grid where: there is a single  $X$  and a single  $O$  in each column and row; vertically the  $X$ 's are connected to the  $O$ 's; horizontally the  $O$ 's are connected to the  $X$ 's; all vertical lines are overcrossings; the grid has torus identifications.

Recently, Sarkar [17] defined several maps between the HFK complexes of two different knots. In particular, he defined chain maps, whose underlying grid diagram maps could be viewed as births, deaths and saddles of a grid cobordism connecting two grid diagrams. Juhász [8] using sutured Floer homology has already shown that  $\widehat{HFK}$  is functorial with respect to smooth decorated cobordisms. This leads one to conjecture the following.

**Conjecture 1.1.** *Grid diagram maps induce maps on  $HFK^-$  that are invariant with respect to marked smooth isotopy classes of surfaces.*

At this point, this conjecture is premature for at least two reasons: (1) the correspondence between smooth surfaces and these sequences is not understood which also means that it is not understood how smooth surface isotopies relate to the yet undefined notion of grid isotopy; (2) little is known about marked smooth isotopy classes. The purpose of this paper is to present the correspondences and definitions needed for the first point. That is, we define grid movies and grid movie isotopies and give the correspondences between, respectively, smooth embedded surfaces in  $\mathbb{R}^4$  and smooth isotopies of these surfaces. We address the marking concerns dealing with the second point in [7].

However, before moving on we give a brief explanation of why the marking correspondences are necessary. It was conjectured that knot Floer homology is only well defined with a particular type of marking. In fact Sarkar in [18] has shown that rotating this marking once around a knot induces a non-trivial automorphism of the knot Floer homology chain complex for most of the 85 prime knots up to nine crossings. Therefore, we eventually need to understand the relation between marked grid movies and smooth marked surfaces embedded in 4-space. However, we will only deal with the unmarked topological aspects in this paper.

In the rest of this section we present a brief introduction to different approaches used to study surfaces smoothly embedded in  $\mathbb{R}^4$  and their smooth isotopy classes and then give an outline of the paper. However we first note that, even though the inspiration to study grid movies came from HFK, the results in this paper are purely topological in nature and should provide a framework to ask naturality questions of other theories that use grid diagrams.

The main idea is to use a sequence of grid diagrams to represent the embedded surface and then study the composition of maps induced on the knot Floer chain complex. Fortunately, studying surfaces using a collection of knot diagrams is not a new idea. Fox studied surfaces using level sets in [6] and Roseman used knotted surface diagrams in [16] and Carter and Saito in [3] combined these two approaches and used a continuous collection of knot diagrams to characterize embedded surfaces in  $\mathbb{R}^4$  up to ambient isotopy. Finally, in [2], Carter, Rieger and Saito reduced the amount of data needed to represent (and distinguish the isotopy type of) smooth embedded surfaces in four space to a specific finite sequence of classical knot diagrams and moves on these sequences. It is this ‘second’ movie move theorem, presented in [2], that will play the dominant role in the current paper since our goal is to present these surfaces as combinatorial objects in terms of grid diagrams. To do this we will need to understand how the combinatorial structure presented in [2] coincides with grid diagrams. Furthermore, since this last theory uses elements of all the previous theories, we will briefly review all of the theories mentioned above presently.

To understand the constructions of these various theories we will begin by reviewing the well known procedure for studying knots in  $\mathbb{R}^3$  (or  $S^3$ ) using knot diagrams. A *knot*  $K = K(S^1)$  is a particular embedding of  $S^1$  in an ambient space  $K: S^1 \rightarrow \mathbb{R}^3$ . A *knot diagram*  $D_p(K)$  is a projection of  $K$  to a plane,  $p: K \rightarrow \mathbb{R}^2$  along with some extra data that keeps track of which strand is the overcrossing at each singular point of the projection. To understand knots up to ambient isotopy the Reidemeister moves are required. Specifically, they account for changes that can occur to the singular set of the projection  $p$  in an ambient isotopy of the embedding space  $\mathbb{R}^3$ .

Embedded surfaces in  $\mathbb{R}^4$  can be studied in a similar way. A *knotted surface*  $K$  is an embedding of some (possibly disconnected) underlying closed surface  $F$  in  $\mathbb{R}^4$ ,  $K: F \hookrightarrow \mathbb{R}^4$ . As in knot theory, we will let  $K$  refer either to the embedding  $K: F \hookrightarrow \mathbb{R}^4$  or the image  $K(F)$  of the embedding. A *knotted surface diagram*  $\mathcal{K}: = p(K)$  is the image of the projection of  $K$  to a hyperplane of  $\mathbb{R}^4$ ,  $p: K \rightarrow \mathbb{R}^3$ . Roseman [16] provided a set of seven Reidemeister type moves, now known as Roseman moves, that operate on knotted surface diagrams and characterize ambiently isotopic surfaces. Carter and Saito refined this procedure. In [3] they studied knotted surface diagrams by choosing a Morse function that separates the critical points of the manifold and the singular sets of the projection. From this construction they defined a movie of a knotted surface diagram and supplied the Reidemeister type moves, now known as movie moves, needed to characterize ambiently isotopic surfaces that preserve this extra structure. In order to present the isotopy class of a knotted surface by a finite sequence of classical knot diagrams Carter, Rieger, and Saito introduced a second height function in each still and supplied a refined, and more numerous set of movie moves in [2]. Stated briefly, when studying isotopy of surfaces in four dimensional space there are three options: (1) use the knotted surface diagrams and the seven Roseman moves; (2) use movies and the 15 movie moves; or (3) use movies with a height function in each still and use the refined set of 31 movie moves.

We will use the ‘second’ movie move theorem that requires movies of knotted surface diagrams with an additional fixed height function in each still and the 31 movie moves in this paper. However, in order to understand this theorem it is necessary to understand the original movie move theorem.

**1.1. Original Movie Move Theorem.** Intuitively, one should think of a movie as the collection of inverse images associated to a Morse function (or even a projection)  $f: \mathcal{K} \rightarrow \mathbb{R}$ . This gives an ordered collection of stills  $\{f^{-1}(s)\}_{s \in \mathbb{R}}$ , which one can analyze sequentially much like watching a movie by increasing or decreasing  $s$ .

Carter and Saito used movies of knotted surface diagrams to characterize embedded smooth surfaces in  $\mathbb{R}^4$  up to ambient isotopy. Not surprisingly, they needed to keep track of more data in order to do this. Specifically, they needed to understand the critical points of the manifold as well as the critical points of the double and triple point sets of the projection. In the next few paragraphs, when defining movies of knotted surface diagrams, we follow [3] closely.

**Definition 1.2.** *For a closed surface  $F$  a smooth map  $f: F \rightarrow \mathbb{R}^3$  is generic if it is an immersion except for some isolated points, which are cone points of figure 8’s.*

The image of a generic map satisfies the following conditions. For all  $y \in F$  there exists a neighborhood  $N(y)$  such that  $(N(y), f(F) \cap N(y))$  is diffeomorphic to:  $(B^3, \text{intersection of } i \text{ coordinate planes})$  where  $i = 1, 2, 3$  and  $B^3$  contains the origin; or  $(B^3, \text{the cone on a figure 8 curve})$  where the figure 8 curve is in the boundary of  $N(y)$ .

The  $j$ -tuple set is  $S_j = \{y \in \mathbb{R}^3; : \#f^{-1}(y) = j\}$ . For  $j = 2, 3$  this is the double point and triple point sets respectively. A *branch point* is a point  $y \in \mathbb{R}^3$  such that the intersection of any neighborhood  $N(y)$  with  $f(F)$  contains a cone on a figure 8. Note that the closure of the

double point set  $\overline{S_2}$  contains the branch points and the triple points (see picture on the right of figure 1).

We would like to do more than just denote the double and triple point sets. We will need critical point data of these sets, which means we need to understand the manifold structure of  $\overline{S_2}$ . To do this, note that  $\overline{S_2}$  is the image of a compact 1-dimensional manifold (non-generically) immersed in  $\mathbb{R}^3$ . To define this curve we will use the configuration space

$$\tilde{C}_r = \{(x_1, \dots, x_r) : x_j \in F, x_j \neq x_k \text{ for } j \neq k, f(x_1) = \dots = f(x_r)\}.$$

There is a free action of the permutation group,  $\Sigma_r$  on this space. The  $r$ -*decker* manifold is the associated  $r$ -fold covering space

$$D_r = \tilde{C}_r \times_{\Sigma_r} \{1, 2, \dots, r\}.$$

The quotient  $C_r = \tilde{C}_r / \Sigma_r$  is the  $r$ -*tuple manifold*. There are maps

$$\begin{array}{ll} \tilde{f}_r : D_r \rightarrow F & f_r : C_r \rightarrow \mathbb{R}^3 \\ [(x_1, \dots, x_r); j] \mapsto x_j & [x_1, \dots, x_r] \mapsto f(x_1) \end{array}$$

that make the square (involving the function  $f$ ) in the following diagram commute and satisfy  $f_r(C_r) = \cup_{i=1}^r S_i$  for  $r = 1, 2, 3$ .

$$(1) \quad \begin{array}{ccccc} D_r & \xrightarrow{q} & C_r & & \\ \tilde{f}_r \downarrow & & \downarrow f_r & & \\ F & \hookrightarrow & \mathbb{R}^4 & \xrightarrow{p} & \mathbb{R}^3 \xrightarrow{p_1} \mathbb{R} \\ & \searrow f & & & \end{array}$$

The map  $q : D_r \rightarrow C_r$  is the covering projection. The image of  $f_r$  is the  $r$ -decker set.

To include the branch points in this picture, we will consider them to be in the double point set even though the preimage of a branch point  $y$ , is a single point. With this understanding, branch points are precisely the boundary of the double point set that lie in the interior of  $F$ . For consistency, we include  $f^{-1}(y)$  in the double decker manifold and extend the covering  $q$  to a branched covering.

**Definition 1.3.** For a fixed embedding  $F \hookrightarrow \mathbb{R}^4$  and generic map  $f : F \rightarrow \mathbb{R}^3$  (where  $f := p \circ K$  as in diagram 1) a projection  $p_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$  is a generic Morse function for the knotting if:

- (1)  $p_1 \circ f_r$  has only non-degenerate critical points for all  $r = 1, 2, 3$ ;
- (2) each critical point is at a distinct critical level of  $p_1$ . (Define critical points to include triple points and branch points.)

Throughout this paper we will always choose to draw the Morse height function in the horizontal direction: see figure 1 for an example.

**Definition 1.4.** A smooth movie of the knotted surface  $K$  is a knotted surface diagram  $\mathcal{K}$  with a fixed generic Morse function.

Using Morse theory, each movie (of a knotted surface diagram) gives rise to a knotted surface and each knotted surface gives rise to a movie.

**Definition 1.5.** A complementary coordinate system for an embedding is an orthonormal coordinate system of  $\mathbb{R}^4$ ,  $(v, v_1, v_2, v_3)$  that satisfies:

- (1) the projection orthogonal to  $v$ ,  $p: \mathbb{R}^4 \rightarrow \mathbb{R}^3$  is generic;
- (2) projection of  $p(\mathbb{R}^4)$  onto the vector  $v_1$  is a generic Morse function  $p_1: \mathbb{R}^3 \rightarrow \mathbb{R}$  for  $K$  and  $p \circ K$ .

A complementary coordinate system can be chosen for any knotted surface. To see this, choose any orthonormal system and if the conditions are not satisfied for  $v$  and  $v_1$  perturb  $v$  while remaining in  $v_1^\perp$  to make the projection  $p$  generic. Similarly, if needed, perturb  $v_1$  while remaining in  $v^\perp$  to get a generic Morse function  $p_1$  for  $K$  and  $p \circ K$ . Finally, choose  $v_2 \in \text{span}(v, v_1)^\perp$  and  $v_3 \in \text{span}(v, v_1, v_2)^\perp$ .

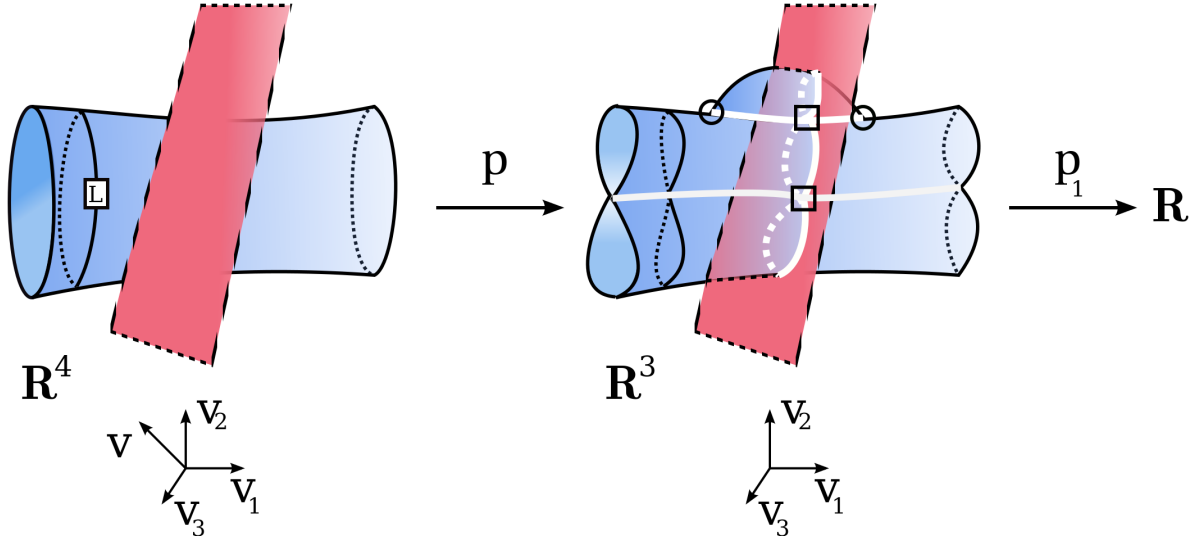


FIGURE 1. Two subsets of a surface  $K$  are tracked through a projection  $p$  and Morse function  $p_1$ . The  $L$  appearing on the left indicates the knot type of that particular level set of  $K$ . White lines represent the double point set  $S_2$ , the squares contain triple points and the circles contain branch points.

Let us pause for a moment to give an intuitive picture of how to think about these constructions. Let  $g: \mathbb{R}^4 \rightarrow \mathbb{R}$  be the projection onto  $v_1$  and for simplicity assume  $g$  is a Morse function for the surface  $F$ . Since [6] people have studied surfaces in  $\mathbb{R}^4$  using level sets. Denote these level sets of  $\mathbb{R}^4$  by  $\mathbb{R}_t^3 := g^{-1}(t)$ . For all but finitely many values of  $t \in \mathbb{R}$ , the intersection of the embedded surface  $K$  with  $\mathbb{R}_t^3$  will be empty or a link. In figure 1 two subsets of a surface  $K$  are shown on the left and their images under  $p$  are shown on the right. The  $L$  appearing in one of the subsets indicates the knot type of that particular level set, which in this example is the unknot (since  $p(L) \cap S_2$  is a single point).

By glancing at the projection of the simple surfaces appearing in figure 1 it is easy to see that knotted surface diagrams can get quite complicated, which makes it difficult to keep the relevant structure in the mind. Since a movie (of a knotted surface diagram) is defined to be the level sets of the projection  $p$  it might seem that we are stuck using knotted surface diagrams. However, it turns out that much of the structure can be understood using the level sets of  $g$  as long as a complementary coordinate system is chosen. A complementary coordinate system allows one to reverse the order of the projections: slice then project rather than project then slice. For example, in the diagram on the left of figure 1 consider the level set  $\mathbb{R}_s^3$  that has the link component  $L$ . With a complimentary coordinate system  $p(L) = p_1^{-1}(s)$ . That is to say, the projection  $p$  simply takes the link  $L$  to a diagram which is the still of the movie. Said another way, the projections of the level sets of  $g$  are diagrams, which happen to be the stills (level sets of  $p_1$ ) corresponding to the movie of  $\mathcal{K}$ . Furthermore, since both  $g$  and  $p_1$  are defined by projection onto  $v_1$  the indexing of the level sets is the same (*i.e.*  $p \circ g^{-1}(s) = p_1^{-1}(s)$ ).

**Definition 1.6.** *Two knotted surfaces  $f_i: F \rightarrow \mathbb{R}^4, i = 0, 1$  are ambiently isotopic if there is an isotopy  $H: \mathbb{R}^4 \times [0, 1] \rightarrow \mathbb{R}^4$  s.t.  $H(x, 0) = x \forall x \in \mathbb{R}^4$  and  $H(f_0(a), 1) = f_1(a) \forall a \in F$ .*

The main theorem of [3] is the following.

**Theorem 1.7.** *(Carter and Saito 1993) Two knotted surface movies represent isotopic knottings if and only if they are related by a finite sequence of movie moves 1-7, 23a, 23b, 25-30 (figure 19) or interchanging the levels of distant critical points.*

The smooth model movie moves 1-7, 23a, 23b, 25-30 (depicted in figure 19) can vary crossing information and can be interpreted as sequences starting at the top or starting at the bottom. These moves also represent moves where each still has been reflected horizontally (or vertically or both).

When studying knots using diagrams, the Reidemeister moves are needed to recover ambient isotopy of knots. However, the full statement is that the diagrams of any two ambient isotopic knots are related by Reidemeister moves *and* planar isotopies. Correspondingly, the theorem above leaves the type of isotopy implied. Two knotted surfaces represented by movies (of knotted surface diagrams) are isotopic if they are related by movie moves, exchanging critical points *and* level preserving isotopies of  $\mathbb{R}^3$  (*i.e.* an isotopy through movies). This point is crucial to understanding how to derive a grid movie move theorem.

**1.2. The Second Movie Move Theorem.** Unfortunately, the first movie move theorem does not give a combinatorial description of a knotted surface. Fortunately, Carter, Rieger and Saito in [2] provided the needed combinatorial structure with their second movie move theorem. We will explain these constructions in the next subsection. Presently, we sketch a brief argument that shows why the structure present in the first movie move theorem is not robust enough to represent smooth movies combinatorially, which also happens to be one of the motivating factors for the second movie move theorem.

Suppose you are given a smoothly embedded surface in four space and then construct a movie of a knotted surface diagram that corresponds to this surface. It is a simple matter to

reconstruct a smooth surface that is isotopic to the original since you have the full set of stills from the movie of the knotted surface diagram (even though only a finite sub-collection of stills is used for depiction purposes). The question we are currently framing is the following, “Is there a general method of selecting a finite sub-collection of these movie stills that will always allow one to reconstruct a smooth surface that is isotopic to the original? If so, how many stills, and which ones, need to be included in this sub-collection?”

One might initially postulate that only one still in between each of the type I, II, and III critical points is needed since all stills between two successive critical points of the Morse function are planar isotopic to one another. Carter and Saito showed that this sub-collection of stills is not sufficient to reconstruct a surface isotopic to the original. They did this using an example of a surface that had stills between two successive critical points that, when viewed sequentially in the Morse time direction, looked like a trefoil rotating in a plane. This example can be found in [2] (section 3.4.1 and corresponding figure 15). Briefly, it is possible for the link diagram to ‘rotate’ (a planar isotopy) in between successive critical points, as the time parameter of the Morse function changes, inducing a ‘twist’ in the surface, which can alter the surface isotopy class. The main point is, given a finite sub-collection of stills with only one still in between critical points of type I, II, and III one would not be guaranteed to reconstruct a surface in the same isotopy class as the original because there is ambiguity in how many times each link diagram is suppose to ‘twist’ in between each successive pair of critical points.

In order to keep track of these ‘twists’ Carter, Rieger and Saito in [2] introduced a second fixed height function. We only briefly describe their constructions and point the interested reader to their paper for more details. Specifically, given a knotted surface  $K$  embedded in four space, they: project to an  $\mathbb{R}^3$  hyperplane to get a knotted surface diagram ( $p: \mathbb{R}^4 \rightarrow \mathbb{R}^3$ ); project onto a retinal plane  $\pi: \mathbb{R}^3 \rightarrow P$  (an affine plane  $P$  that does not intersect the image  $p(K)$ ); lastly they define a height function  $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ .

$$F \xrightarrow{K} \mathbb{R}^4 \xrightarrow{p} \mathbb{R}^3 \xrightarrow{\pi} P \xrightarrow{h} \mathbb{R}$$

They then analyze the singularities, with respect to  $h$ , of the projection onto the retinal plane.

**Definition 1.8.** *The inverse images of the  $t = a$  and  $t = b > a$  levels differ by an elementary string interaction (ESI), with respect to the movie description with a still height function, if only one singular value of the height function  $h$  occurs between  $a$  and  $b$ .*

There are seven basic types of ESI’s, which are represented by the diagrams in the top row and the two diagrams on the bottom left of figure 2. Starting from the top left and reading to the right they correspond to: branch points; maximal or minimal critical points occurring on the interior of a double point arc; isolated triple points (formed by three sheets); birth or death of an unknotted circle (Morse critical point of index 0 or 2); saddle critical point (Morse critical point of index 1); a *cusplike* or a *switchback move*; *camel-back move* or a  $\psi$  move.

In order to obtain a combinatorial description of a knotted surface the ‘multi-local’ moves, which account for exchanging the levels of distant critical points (depicted in the bottom right four diagrams in figure 2), must be included.

Together the 11 elementary string interaction’s in figure 2 make up the full set of elementary string interactions (FESIs). Each FESI has several versions obtained by: reflecting each frame

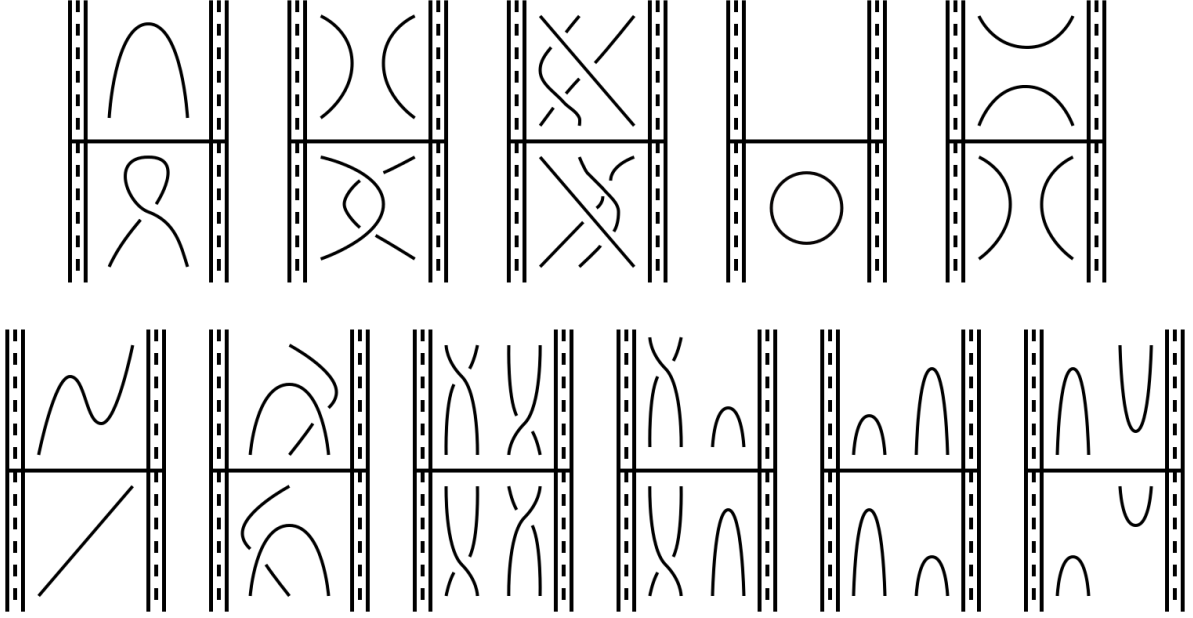


FIGURE 2. The full set of elementary string interactions (FESI's): the five movie move ESI's in the top row; the right four on the bottom row are referred to as multi-local moves. The left two ESI's on the bottom row are the cusp and  $\psi$ - moves.

about the horizontal (or vertical) axis; changing between positive and negative crossings; orienting the strings differently; and reading the film strip from bottom to top rather than top to bottom. Taking into account these symmetries, the FESIs represent the stills that need to be included in the finite sub-collection in order to reproduce a smooth surface of the same isotopy class as the original (see theorem 3.5.4 in [2]). More explicitly, to generate a sufficient sub-collection of stills to reconstruct a surface of the same isotopy class: using the time parameter, order the stills that contain either a Morse critical point of type I, II, III or contain a singular point of the height function  $h$ ; and then select a single still in between each of these critical points. This ensures that each subsequent still in this sub-collection differs from the previous one by a FESI (and possibly a level preserving isotopy of  $\mathbb{R}^3$  that respects  $h$ ).

One of the many results in [2] is the following theorem.

**Theorem 1.9.** (Carter, Rieger, Saito 1997) *Two smooth movies with height functions on each still represent isotopic knotted surfaces if and only if one can be obtained by the other by a finite sequence of the 31 smooth movie moves with height functions fixed in each still.*

In this paper we prove the following grid diagram analogue of the above theorem.

**Theorem 1.10.** *Two grid movies represent isotopic knotted surfaces if and only if one can be obtained from the other by a finite sequence of grid movie moves.*



These moves are depicted in figure 19.

In order to generalize this movie move theorem to a grid diagram analogue a few things need to be understood: (1) how to gridify a movie; (2) what a grid planar isotopy is (so that we may define a level preserving grid isotopy of a grid movie); and (3) how to represent a movie (of a knotted surface diagram), up to isotopy, by a finite collection of classical knot diagrams.

The second movie move theorem clearly deals with the third point. The main idea to deal with the first point is to gridify each still, while preserving the planar isotopy class, in a combinatorial description of a knotted surface.

Both Cromwell [4] and Dynnikov [5] have provided algorithms that take a smooth knot diagram to a grid diagram. In addition they have shown

**Theorem 1.11.** (*Cromwell (1995), Dynnikov (2006)*) *Any two grid diagram presentations of a link  $L \subset S^3$  are related by a finite sequence of commutations, stabilizations and destabilizations.*

We will define these three moves in the next section. At present, note that this theorem characterizes grid presentations of links up to ambient isotopy of  $\mathbb{R}^3$  or  $S^3$ . However, it does not characterize grid presentations of links up to planar isotopy, which is needed for the problem at hand. Although we can not directly use Cromwell and Dynnikov's theorem, it will help us define a correct notion of grid planar isotopy.

The rest of this paper is organized as follows. We give a definition of grid planar isotopy. We show that grid planar isotopy coincides with the smooth definition. That is, the set of equivalence classes of smoothly planar isotopic diagrams is in one to one correspondence with the set of equivalence classes of grid planar isotopic diagrams. In order to show this, we present a different planar grid algorithm that takes a smooth link and returns a grid link. We then define grid movies and grid movie moves. Following this we describe how to obtain a grid movie from a smooth movie with a fixed height function in each still. In order to do this we will slightly modify our grid planar algorithm to account for a fixed height function. We will then prove theorem 1.10. Finally, since we will be concerned with surfaces with boundary components, we generalize the definitions and constructions to handle this more general setting.

## 2. GRID PLANAR ISOTOPY

As mentioned in the introduction, both Cromwell [4] and Dynnikov [5] have given independent proofs of theorem 1.11. Cromwell used Markov moves and Dynnikov used Reidemeister moves.

Theorem 1.11 shows that commutations, stabilizations and destabilizations of a grid diagram (which will be defined shortly) encode both planar isotopy information and the Reidemeister moves. Therefore, since we seek a grid equivalent of planar isotopy, we will remove the types of commutations, stabilizations and destabilizations that encode Reidemeister moves. It turns out that we do need to allow a specific combination of two Reidemeister 2 moves, which

we call a transfer move, to fully capture the right notion of grid planar isotopy. We will deal with stabilizations and destabilizations together, enumerate the types of commutations that can occur and then discuss transfer moves.

First, since a grid diagram has torus identifications we will choose a fundamental domain by specifying a vertical arc and horizontal arc, which we label  $\alpha_0$  and  $\beta_0$  correspondingly. The intersection point  $\alpha_0 \cap \beta_0$  will always be in the lower left corner of our diagrams. We index the other vertical arcs sequentially left to right and horizontal arcs from bottom to top. Columns are indicated by the left vertical arc of the column and rows are indicated by the lower horizontal arc. For a grid diagram of index  $n$  there are columns and rows indexed by  $0, 1, \dots, n-1$ .

A *stabilization* of a grid diagram takes a grid of index  $n$  to a grid of index  $n+1$  by the following procedure: choose a row (column) and split it into two adjacent rows (columns); add a new column (row) in the grid diagram; place an  $X$  and an  $O$  adjacent to one another such that they are in the two modified rows (columns) and the new column (row). A *destabilization* is the opposite of this procedure: delete an  $O$  and  $X$  that are in adjacent rows (columns); delete the horizontal (vertical) arc that separate them; delete one of the vertical (horizontal) arcs on either side; contract the two rows and two columns. Examples of stabilizations and destabilizations are shown in figure 3.

The stabilizations and destabilizations either add another crossing (Reidemeister 1) or they don't, in which case they just add a kink. See figure 3 for examples. Define a de/stabilization move to be a *kink de/stabilization* if it does not add an intersection point and define it to be an *R1 de/stabilization* otherwise see figure 3.

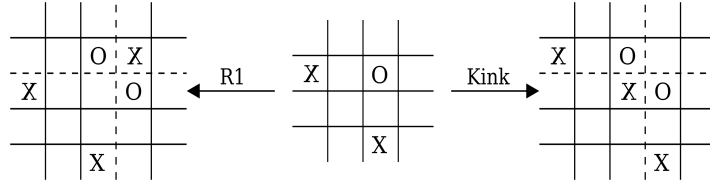


FIGURE 3. An R1-stabilization (left) and a kink stabilization (right). The dotted lines are the lines added to a grid diagram. Inverting this process yields examples of the two types of destabilizations.

A *commutation* move is an interchange of two adjacent rows (or columns). Not all interchanges are allowed. In order to determine if a commutation can occur: label the  $X$ 's and  $O$ 's by  $X_1, O_1, X_2, O_2$  where  $X_i, O_i$  are in the same column or row; each of these has row and column coordinates; for a commutation of two columns (rows) order these by the row (column) coordinate; if two share the same row (column) number or if neither  $X_1, O_1$  nor  $X_2, O_2$  are sequential in this ordering the commutation is obstructed. An  $\alpha_i$  *commutation* is a commutation between two rows that share the horizontal arc  $\alpha_j$ . Similarly, a  $\beta_j$  *commutation* is a commutation between two columns that share the vertical arc  $\beta_i$ .

There are two types of commutations that can occur: *un-nested* where both pairs  $X_i, O_i$  are sequential; and *nested* where exactly one pair of  $X_i, O_i$  are sequential. Even though  $\alpha_0$

and  $\beta_0$  commutations are defined we will not use them since pulling arcs around the back can represent Reidemeister one moves (see the top left diagram in figure 4). An un-nested commutation is always a grid planar isotopy. A nested commutation can represent three things: (1) a Reidemeister 2 move ( $R2$ ) that changes the number of intersections by  $\pm 2$ ; (2) Reidemeister 3 move(s) ( $R3$ ) in which the separated pair of points (representing an arc of the knot) moves an intersection point over another intersection point; (3) a grid planar isotopy move ( $P$ ) in which the separated pair of points moves an intersection along an arc of the knot. See figure 4 for examples. A word of caution: many times, when dealing with grid diagrams, the knot is not drawn, which means care is needed when trying to do a grid planar isotopy since the only difference between the  $R3$  move and a  $P$  move is the existence of the horizontal arc separating both  $X, O$  pairs. Furthermore, the only difference between an  $R2$  move and a  $P$  move is the location of one  $O$  (or  $X$ ). Compare the  $R2$  and  $R3$  moves against the  $P$  move in figure 4.

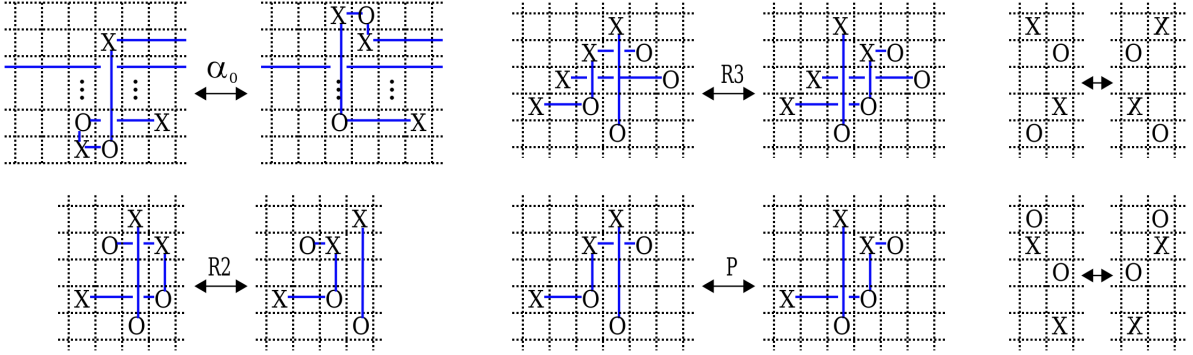


FIGURE 4. Clockwise from top left: an  $\alpha_0$  commutation that represents an  $R_1$  move,  $R_3$  move, nested pair of arcs, un-nested pair of arcs, a  $P$  move and an  $R2$  move.

Suppose that we defined a grid planar isotopy to be a composition of kink destabilizations, kink stabilizations, un-nested commutations and  $P$  moves that are neither  $\alpha_0$  nor  $\beta_0$  commutations (all the non-Reidemeister moves). Then, unfortunately, there would be no grid planar isotopy representing the process of smoothly isotoping an intersection point around a bend in an arc (see figure 5). Notice that the arc that is being isotoped changes its direction in the plane by  $180^\circ$ . This together with the fact that commutations and kink de/stabilizations will never reverse the orientations of arcs means that we cannot represent all planar isotopies with the non-Reidemeister grid moves. To fix this issue we need to add one more move which is a very specific composition of two Reidemeister II moves called a transfer move depicted in figure 5.

**Definition 2.1.** A transfer move is a composition of two  $R2$  moves, one horizontal and one vertical, between two arcs involved in an intersection point  $p$ . The first  $R2$  move creates two new intersection points of opposite sign between these two arcs. The second  $R2$  move removes  $p$  and the newly created intersection point with sign opposite of  $p$ .

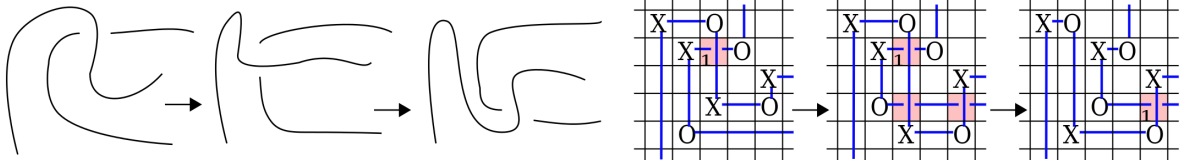


FIGURE 5. Left: an isotopy that moves the double point smoothly around a corner. Right: the corresponding transfer move applied to a grid diagram representing the initial smooth diagram.

In more detail, the first  $R1$  move either pulls the top strand over the bottom strand or the bottom strand under the top strand, either horizontally or vertically, creating two more intersection points with opposite sign. The second commutation removes the original intersection point and the intersection with the opposite sign as the original, leaving only one of the new intersection points. If the initial intersection point had an  $X$  above and an  $O$  below it then the final intersection point will have an  $O$  above and an  $X$  below (and vice versa).

**Definition 2.2.** A grid planar isotopy is a finite composition of kink (de)stabilizations, transfer moves, un-nested commutations, and  $P$  moves that are neither  $\alpha_0$  nor  $\beta_0$  commutations.

### 3. PLANAR GRID ALGORITHM

The goal of this section is to show that our definition of grid planar isotopy coincides with smooth planar isotopy. Specifically, we present an algorithm for transforming a planar diagram of a knot into a grid diagram and show that it is well defined up to grid planar isotopy. Then, using the structure that the planar grid algorithm provides, we show that smooth planar isotopy classes are in one to one correspondence with grid planar isotopy classes.

We begin with some definitions needed for the algorithm and proofs within this section.

**Definition 3.1.** A grid arc is a closed connected segment of a grid diagram of a knot consisting solely of vertical and horizontal line segments meeting at right angles that connect two points.

**Definition 3.2.** An arcword of a grid arc is some ordering of  $L$ 's and  $R$ 's or possibly the single letter  $I$ . Each  $L$  and  $R$  represents a left turn and right turn respectively. To create an arcword, start at one of the endpoints of the grid arc and travel along the grid arc until the other end point is reached notating each turn that is passed through. If no turns are passed then the arcword is  $I$ .

In this terminology a kink is either a  $LR$  or  $RL$  combination. Many times  $LR$  and  $RL$  combinations can be removed or added by a destabilization and stabilization respectively.

**Definition 3.3.** A reduced arcword is an arcword where all possible  $LR$  and  $RL$  strings have been removed by a grid planar destabilization. If there are no  $L$ 's or  $R$ 's remaining represent this by the arcword  $I$ .

**Definition 3.4.** A net turn is an arcword where all  $LR, RL, LLLL$ , and  $RRRR$  terms are algebraically set to the identity  $I$ .

For example, the section of the two blue grid arcs in figure 6 (assuming that the orientation is left to right) would induce a word  $I$ , and the word  $RLLRRL$  representing the six turns with direction. Note that both of these arcs have the same reduced arcword of  $I$  (and therefore the

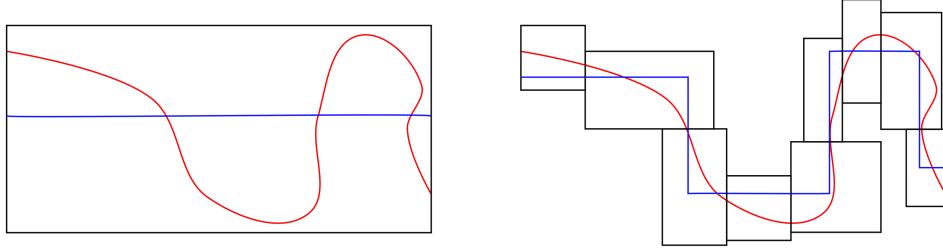


FIGURE 6. Two possible rectangle choices for the original smooth arc.

same net zero turn). This is true in general: any two sets of rectangles chosen for the same arc will have the same reduced arcword.

**Definition 3.5.** An intersection parallelogram, for a knot diagram  $D$ , is a parallelogram  $P$  with horizontal sides such that:  $P$  contains a single double point and exactly two arcs in its interior; one arc intersects each of the horizontal sides exactly once, call this the vertical arc; the other arc intersects each of the other two parallel sides exactly once, call this the horizontal arc; the arcs do not intersect the corners of  $P$ .

**Definition 3.6.** An intersection square  $R$  is a square that contains a single double point and exactly two arcs in its interior, an intersection parallelogram  $P$ , two trapezoids  $T_L, T_R$  and four rectangles according to the following rules (see figure 7 for examples):

- (1)  $\partial R \cap \partial P = \emptyset$
- (2) The trapezoids  $T_L, T_R$  each have at least two right angles;  $T_L$  ( $T_R$ ) has a side that is contained in the left (right) non-horizontal side of  $P$ ; this side contains a single intersection point with the horizontal arc contained in  $P$ .
- (3) Each of the four rectangles shares a different corner of  $R$ .
- (4) The boundary of one rectangle intersects the upper horizontal side of  $P$  and the boundary of another rectangle intersects the lower horizontal side of  $P$ .
- (5) The boundary of the other two rectangles intersect a side of one of the trapezoids.
- (6) The boundary of each trapezoid (except for  $P$ ) has exactly two intersection points with a single arc.

**Definition 3.7.** The planar grid algorithm applied to a diagram  $D$  associated to a projection  $\pi$  of knot  $K$  is enumerated below.

- (1) Draw an intersection square around each double point.
- (2) Number the intersection squares and letter all of the arcs that connect these squares together. The lettered arcs will have endpoints lying in the boundary of intersection squares and will be disjoint from the interior of all intersection squares.

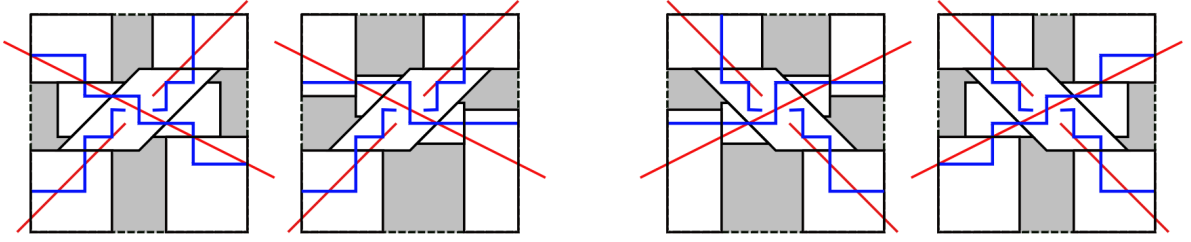


FIGURE 7. Four possible choices of parallelogram and trapezoids involved in an intersection square. The shaded regions are exterior to all drawn shapes, straight lines are the original arcs, and the arcs with  $90^\circ$  turns are the grid arcs.

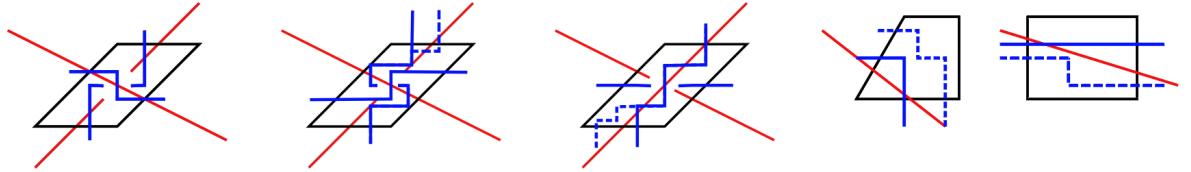


FIGURE 8. Three choices for replacing the original smooth arcs, represented as straight lines, with grid arcs for an intersection point. The two choices of travel through the intersection point when the horizontal arc is also the overarc are shown in the left two diagrams. Dotted lines are other possible grid choices.

- (3) Approximate each lettered arc by drawing rectangles, with vertical and horizontal sides, such that the arc intersects exactly two distinct edges transversally, each exactly once.
- (4) Replace the arcs in the quadrilaterals by grid arcs. If the quadrilateral is an intersection parallelogram then replace the arcs as shown in figure 8. For trapezoids without double points replace the grid arcs with:
  - (a) a grid arc that has net turn  $0^\circ$  if the arc intersects two edges that are not connected by a vertex;
  - (b) a grid arc that has net turn of  $\pm 90^\circ$  if the arc intersects two edges that are connected by a vertex.
- (5) Delete all the rectangles used for approximation and the original smooth diagram.
- (6) Make sure that no two straight grid arcs (in the whole diagram) are collinear. Make small translations of lines as needed to achieve this.
- (7) Choose a minimal set of horizontal and vertical lines that separate each horizontal and vertical grid arc from all others.
- (8) Make these separating lines uniformly spaced and starting at one corner place an  $X$  then choose a direction to travel along the arc and alternately label the corners  $X$  then  $O$ . Continue until all corners are labeled. If the diagram is of a link this process will need to be repeated for each link component. Note that this will induce an orientation on each component of the link. Therefore, if the diagram is already oriented then the

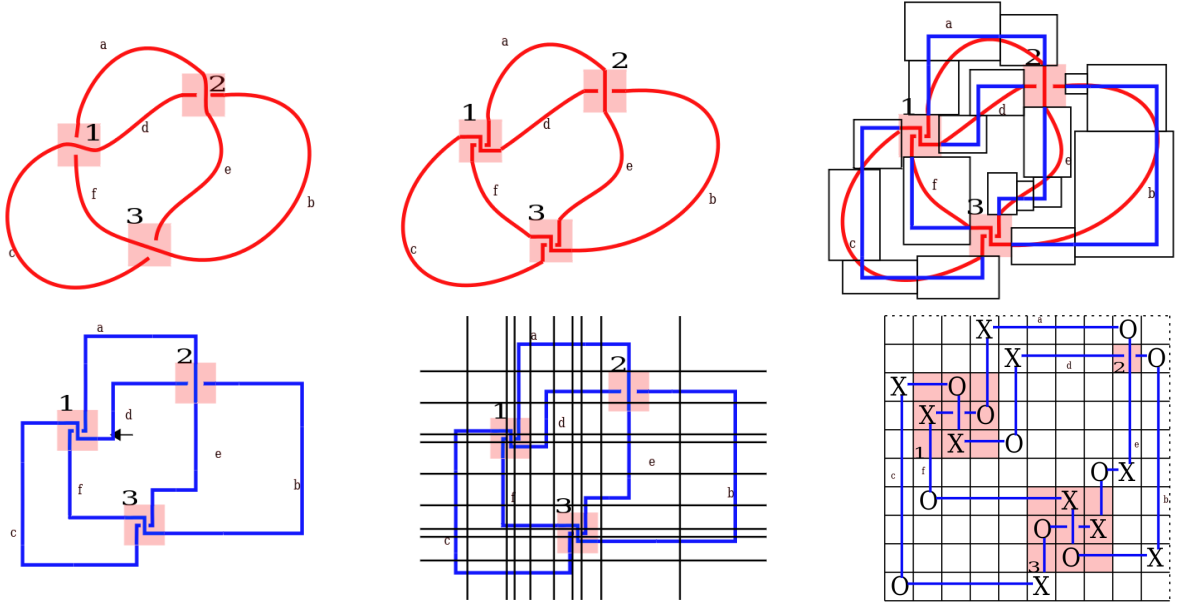


FIGURE 9. The algorithm applied to the trefoil. The stages are as follows: top left and middle diagrams, steps 1, 2 and 3; top right, step 4 and 5; bottom left, step 6 (the arrow indicates the vertical line was moved to satisfy non-colinear condition in step 4); bottom middle, step 7; bottom right, steps 8 and 9.

initial  $X$  placement (of each link component) should be chosen to coincide with the existing orientation.

- (9) Add one more vertical and one more horizontal separating line and make torus identifications to get to a toroidal grid diagram.

**Lemma 3.8.** *There exists an intersection parallelogram for every double point of a diagram of a knot.*

*Proof.* We will represent these arcs by graphs of a function. If either arc involved in a double point has an infinite or zero slope then simply draw a rectangle with horizontal and vertical sides to satisfy the conditions of an intersection parallelogram.

Next, suppose that neither arc has an infinite or zero slope. In almost all cases the conditions of an intersection parallelogram can again be satisfied with a rectangle with horizontal sides. However, a rectangle with horizontal sides will not suffice in general. For example, the two arcs comprising a double point may have slopes of 1 and -1 in the diagram of the knot. To handle the general case let  $p$  be an intersection point in some projection diagram and linear approximate the two arcs  $f_1, f_2$  involved in the crossing by  $l_1, l_2$ . Then we have  $f_i = l_i + R_i$  where  $R_i$  is the remainder, or error term. If  $|f_i''(x)| \leq M_i$  on some interval  $(p-r, p+r)$  then  $M_i \frac{r^2}{2} \geq |R_i(x)|$ . Let  $M = \max\{M_1, M_2\}$ ,  $H$  the vertical distance between the point  $p$  and the horizontal lines, and  $L$  the horizontal distance between the other set of parallel lines (see figure 10). Suppose, that we chose to make the non-horizontal set of lines parallel to the linear

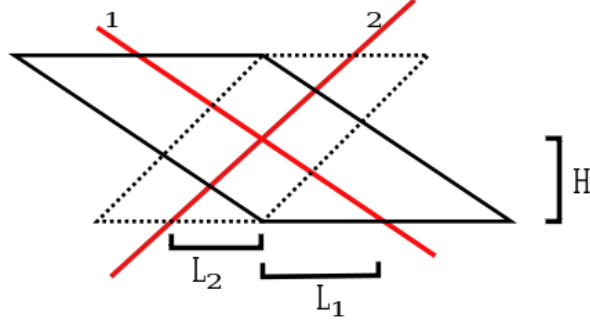


FIGURE 10. Example intersection parallelogram.

approximation of  $f_i, l_i$ . Then what we need to show is the following

$$r \geq H \geq M \frac{r^2}{2} \quad \text{and} \quad r \geq L_i \geq M \frac{r^2}{2}.$$

The left inequalities, of both equations, demand that the parallelogram be inside the region where the inequality for the error term holds and the right inequalities guarantee that the curves have the desired properties (entering and exiting in the desired manner). Using  $L_i = \frac{H}{B_i}$ , where  $B_i$  is the slope of the line  $l_i$ , we see that our inequalities are in terms of  $M, B_i, H, r$ . Since the quantities  $M$  and  $B_i$  are fixed, we only need to choose  $r$  and  $B$  such that the constraints are satisfied. Some algebra shows that by letting  $M = c_M 10^m$  and  $B_i = c_i 10^{b_i}$  where  $c_M, c_i \in (1, 10)$  and choosing  $r = 10^{(-m-2b_i-2k)}$  and  $H = \frac{c_i}{2} 10^{(-m-3b_i-3k)}$  where  $k$  is a positive integer chosen to make  $(-2b_i - k) < 0$  satisfies the inequalities.

Hence there also exists a choice of parallelogram for all non-zero and non-infinite slopes  $\square$

**Lemma 3.9.** *Let  $G(A)$  and  $G'(B)$  be any two oriented grid planar arcs without double points satisfying the conditions below, where  $A$  and  $B$  are smooth planar arcs without double points and  $G(A)$  and  $G'(B)$  represent two possibly different applications of the planar grid algorithm on these arcs.*

- (1) *The endpoints coincide in a superposition of the two grid diagrams.*
- (2) *The initial (resp. final) segments of  $A$  and  $B$  agree.*
- (3) *The arcs have the same net turn.*

*Then  $G(A)$  and  $G'(B)$  are grid planar isotopic arcs.*

*Proof.* This proof follows essentially the same methods used to show the PL-Schoenflies theorem (see [1] page 20-21). Let  $G(A)$  and  $G(B)$  be given as above. We need to give a procedure to take the grid arc  $G(A)$  to the grid arc  $G'(B)$ . Since these grid arcs are in two different grid diagrams there may be complications in giving a planar grid isotopy taking one grid arc to the other. Specifically, the other arcs may initially obstruct a desired grid planar isotopy. An example of this sort of obstruction is in the first diagram in figure 11.

Initially we assume that  $G(A)$  and  $G'(B)$  don't intersect other grid arcs in the superposition.

Let  $S$  be the superposition of  $G(A)$  and  $G'(B)$  such that the endpoints agree.



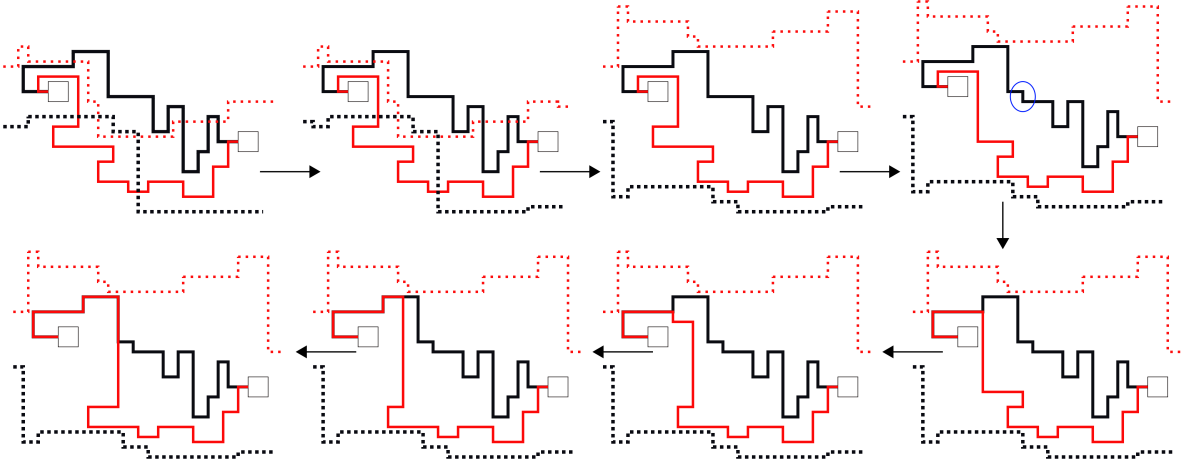


FIGURE 11. Two grid diagrams superimposed, one with red (or lighter gray) arcs and one with black arcs. The solid arcs are the two arcs we want to resolve. The dotted arcs are other arcs that might initially obstruct desired grid planar isotopies. Note that any apparent double points are an artifact of superimposing the two diagrams.

Algorithm for resolving  $G(A)$  and  $G'(B)$  starting from the simple situation depicted in the third diagram on the top row of figure 11:

- (1) Stabilize or destabilize arcs  $G(A)$  and  $G'(B)$  until they have the exact same number of turns (corners). This can be done since  $net(G(A)) = net(G'(B))$ . (The circle in the top right diagram denotes where a stabilization is performed.)
- (2) Perform commutations to move the next turn in each diagram to correspond (bottom right diagram).
- (3) If the grid arcs turn in the same direction do nothing. If they do not, then perform a destabilization followed by a stabilization that changes the direction of the turn (middle bottom diagrams).
- (4) Repeat steps 3 and 4 until  $G(A) = G'(B)$ .

Now assume that there are other arcs intersecting  $G(A)$  or  $G'(B)$  in the superposition. Note that the condition that neither  $G(A)$  nor  $G'(B)$  has double points on it means that all of the intersections in the superposition occur between an arc from each still. Furthermore, this means that the oriented intersection number must be zero between any two arcs.

To planar grid isotope these arcs out of the way, consider two intersection points of opposite sign (connected by an arc with no other double points on it). Stabilize so that the intersection point satisfies the properties of the lemma and then apply the algorithm just mentioned to make these two sub arcs agree. Next, apply commutations to one diagram so that the two sub arcs that were in agreement no longer coincide. Repeat this procedure for all pairs of intersection points.

Hence, we can grid isotope the obstructing arcs out of the way and then apply the procedure above to make  $G(A)$  and  $G'(B)$  coincide through a planar grid isotopy.  $\square$

**Proposition 3.10.** *The output of the planar grid algorithm is well defined. That is, for a given diagram  $D$  two choices of rectangles will lead to grid diagrams of the same grid planar isotopy class.*

*Proof.* Let a diagram  $D$  be given and suppose that the numbering of intersections and lettering of the arcs is the same in both applications 1 and 2 of the grid algorithm arriving at grid diagrams  $G_1(D)$  and  $G_2(D)$ .

First, suppose that the intersection squares are chosen the same in both algorithms. Then the only possible variation is in the choice of covering rectangles used to get a grid arc for each of the arcs connecting the intersection rectangles. Lemma 3.9, applied individually to each of these grid arcs, states that any differences in  $G_1(D)$  and  $G_2(D)$  can be rectified by grid planar isotopies. There is no difficulty since each resolved grid arc does not obstruct the resolution of any other grid arc. Therefore, we only need to consider how different choices of intersection rectangles and parallelograms and trapezoids affect grid diagrams  $G_1(D)$  and  $G_2(D)$ .

First, keep the choice of intersection rectangle fixed and consider the changes to  $G_1(D)$  and  $G_2(D)$  that result from different choices of intersection parallelograms and trapezoids. There are three choices that must be made: (1) which linear approximation to have parallel to the non-horizontal sides of the parallelogram; (2) deciding whether the horizontal smooth arc is going to intersect the trapezoids  $T_{iL}, T_{iR}$  in the vertical or horizontal edges; (3) choice of orientation of the grid arc through the intersection point when the horizontal grid arc is the overarc. The two left diagrams in figure 8 give examples of choice 3; choices 1 and 2 are depicted in figure 7. We will show that all of these choices lead to grid diagrams which are grid planar isotopic. We will take care of choice 3 and then choices 1 and 2 together.

The grid diagrams resulting from the two choices of orientation through the intersection point (choice 3) are related by a transfer move, kink stabilizations and kink destabilizations, whose composition is a grid planar isotopy. The sequence of moves is in figure 12.

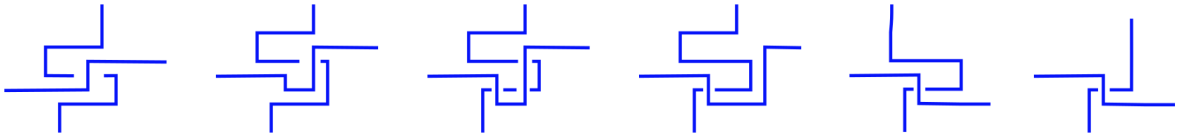


FIGURE 12. A sequence of grid planar isotopies taking one choice of grid arcs in an intersection parallelogram to the other choice. In order from left to right: a stabilization; the transfer move begins with an R2 move; transfer move ends with an R2 move; destabilizations to the last diagram.

If the orientations through the double point do not agree then apply the sequence of moves (or the reverse) in figure 12 so that both grid arcs have the same orientation through the double point. Next apply lemma 3.9 pairwise to the corresponding eight grid arcs (in two

diagrams) with one end point the double point and the other end point being the intersection point of an arc and the intersection square. This means that the choices involved amounted to a grid planar isotopy. As an example, figure 7 depicts the four choices coming from choices 1 and 2. It can be seen that the four choices of grid arcs do indeed only differ by a sequence of kink destabilizations which is a grid planar isotopy.

Hence, we only need to show that different choices of intersection squares also induce differences that are resolved by grid planar isotopies. To this end, we will show that the difference between two choices of intersection squares where one square sits in the interior of the other square gives rise to differences in  $G_1(D)$  and  $G_2(D)$  that are resolved by a grid planar isotopy. Then given two choices of intersection squares say  $E$  and  $F$  we only need to choose a square that sits inside both of them (or contains both of them) to show that this result holds for all choices of intersection squares. Given two intersection squares  $S_1, S_2$  it is clear one can always choose a third intersection square  $S$  such that  $S_2 \supset S \subset S_1$  since double points must reside in the interior of intersection squares.

Suppose that intersection square  $E$  is larger than  $F$  and contains  $F$  in its interior (see figure 13). We have shown that choices of shapes on the interior of intersection squares don't matter

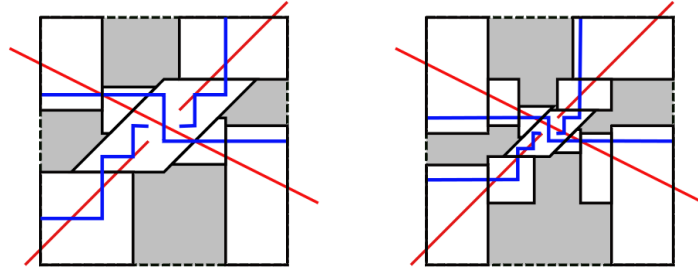


FIGURE 13. Large intersection square  $E$  on the left and intersection square  $F$  on the right.

up to grid planar isotopy. So make some choices for  $E$  and make some choices for  $F$ . In the first case  $G_1(E)$  we simply apply the normal algorithm to get some grid arcs representing the part of the diagram contained in  $E$ . In the second case apply the grid algorithm to the area  $F$  and then choose four more rectangles to cover the arcs connecting the boundaries of  $F$  and  $E$ . By an application of lemma 3.9, each of the four grid arcs in  $E$  (with endpoint the double point and some intersection point between the arc and boundary of  $E$ ) is grid planar isotopic to the corresponding grid arc in  $F$  adjoin the four covering rectangles.

Hence, the output of the planar grid algorithm is well defined.  $\square$

**Theorem 3.11.** *Smooth planar isotopy classes are in one to one correspondence with grid planar isotopy classes. That is two smooth diagrams  $A$  and  $B$  are smooth planar isotopic iff  $G(A)$  and  $G(B)$  are grid planar isotopic.*

*Proof.* The only if direction is fairly straightforward: smooth the corners and observe that since each grid planar isotopy is the initial and final state of a smooth planar isotopy then the

initial and final states of a composition of finitely many grid planar isotopy will also be the initial and final states of a smooth planar isotopy. From the definition of grid planar isotopy it follows that there exists a smooth planar isotopy connecting the beginning and final state.

To see the other direction, let  $D, D'$  be planar isotopic diagrams, let  $G(D)$  and  $G(D')$  be grid diagrams obtained from applying the planar grid algorithm, and let  $h: \mathbb{R}^2 \times [0, 1] \rightarrow \mathbb{R}^2$  be an isotopy with  $h(D, 0) = D$  and  $h(D, 1) = D'$ . We need to show that  $G(D')$  is grid planar isotopic to  $G(D)$ .

There were choices of quadrilaterals,  $Q_0$ , in getting  $G(D)$ . Fix these choices of quadrilaterals and let  $d_1 \in [0, 1]$  be the first value that the quadrilaterals do not satisfy the properties set forth in the grid planar algorithm with respect to the arcs in  $h(D, d_1)$ . We will enumerate the ways in which these quadrilaterals can fail shortly.

At  $d_1 - \epsilon$  choose different quadrilaterals,  $Q_1$  that satisfy the properties of the grid planar algorithm with respect to the arcs in  $h(D, t)$  for  $t \in [d_1 - \epsilon, d_1]$ . Typically, only a few quadrilaterals will be changed. However, if desired or needed, every single quadrilateral could be chosen differently so that  $Q_0 \cap Q_1 = \emptyset$ . The fact that the output of the planar grid algorithm is well defined up to grid planar isotopy (proposition 3.10) means that the different choice of quadrilaterals will lead to grid planar isotopic diagrams.

Repeating the above procedure will yield values  $d_i$  for  $i = 1, \dots, n$  such that the choice of quadrilaterals  $Q_i$  satisfy the needed properties with respect to the arcs in  $h(D, t)$  for  $t \in [d_i - \epsilon, d_{i+1} - \epsilon]$ . Let  $H_i$  be the grid planar isotopy induced by the different choice of quadrilaterals at  $d_i - \epsilon$ . Suppose that  $d_n = 1$  and choose  $Q_n$  to correspond to the choice of quadrilaterals used to get  $G(D')$ . Then  $H := H_n \circ H_{n-1} \circ \dots \circ H_1$  is a grid planar isotopy from  $G(D)$  to  $G(D')$ .

A quadrilateral that fails to satisfy the needed properties required in the grid planar algorithm either involves an intersection parallelogram or not. If it does not involve an intersection parallelogram, then either the smooth arc isotopes around a corner of a trapezoid or it isotopes across a side of a trapezoid (much like a finger move). Five different examples of these two types of isotopies are shown in figure 14. The top two left diagrams have a reduced arcword of  $I$ . The top right has a reduced arcword of  $R$  or  $L$  depending on orientation. In general, many more kinks may be introduced during the isotopy and the arcword may become quite long even though the reduced arcword will not change as can be seen in the diagrams. Many times a choice of larger trapezoid will account for the isotopy. A couple of examples of these can be found in the bottom two finger moves in figure 14.

To deal with an isotopy that requires a change in the intersection square, use the same idea used to prove that the algorithm was well defined. Namely, any time the intersection square is about to fail the conditions set forth in the algorithm (*e.g.* moves across the edge of a parallelogram, etc.) change the intersection square to a small square. Then choose a larger square containing the small square such that this new intersection square satisfies the conditions of the planar grid algorithm with respect to the arcs in  $h(d, t)$  for  $t \in [d_i - \epsilon, d_i]$ . Again these choices of intersection squares only induce a grid planar isotopy difference between the two diagrams by appealing to proposition 3.10.  $\square$

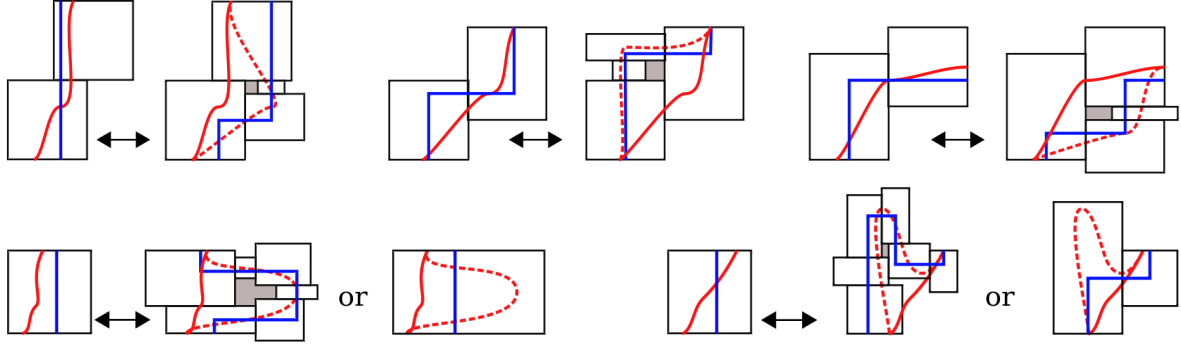


FIGURE 14. Five different examples of failing quadrilaterals that are not intersection parallelograms. The diagrams in the top row isotope a curve past a corner and the diagrams in the second row show two different resolutions for isotoping a curve by a finger move across an edge of a rectangle. The darker lines are the grid diagram arcs, the lighter solid line is the original curve and the dotted lines are the isotoped curve. Grey shaded regions are exterior to all rectangles.

#### 4. GRID MOVIES AND GRID MOVIE MOVES

In this section we prove our main result. In order to do this we first use the results of the last section to define grid movies and grid movie isotopies. We then discuss: compatibility of the combinatorial structures coming from movies and grid diagrams; how smooth movies and smooth movie isotopies are represented by grid movies and grid movie isotopies; model moves; and define grid movie moves.

To define grid movies we first need to define moves to grid diagrams corresponding to type I critical points: births, deaths and saddles (see figure 15).

**Grid birth move:** A *grid birth move* takes a grid diagram  $G$  to  $G'$  with respective grid indices  $n$  and  $n + 1$  by adding a row and column and places an  $O$  and an  $X$  in the square that are both in the new row and column.

**Grid death move:** A *grid death move* is the reverse of a birth: it takes a grid diagram  $G'$  to  $G$  with grid indices  $n + 1$  and  $n$  by removing an  $O$  and  $X$  that are in the same row and column and deletes the row and column.

**Grid saddle move:** A *saddle move* takes a grid diagram  $G$  to  $G'$  with the same grid index by exchanging the places of two  $O$ 's such that one  $O$  is directly left and down from the other  $O$  to the configuration where one  $O$  is directly down and right of the other. This move cannot change the number of double points.

**Definition 4.1.** A grid movie is a finite ordered sequence of grid diagrams such that any two sequential grid diagrams are related by a single grid move from the following list: birth, death, saddle, commutation, stabilization, destabilization and the identity.

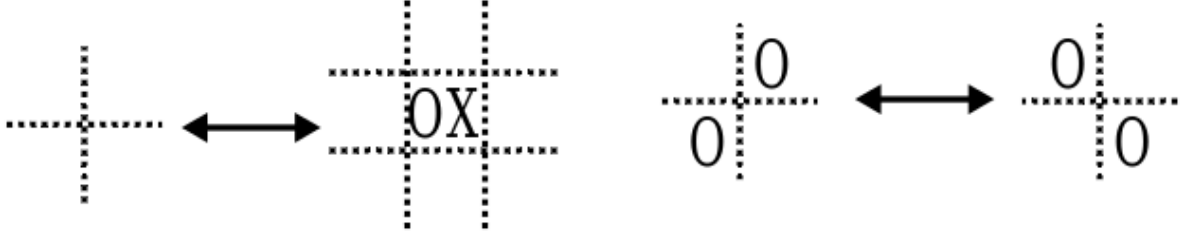


FIGURE 15. Type I critical grid moves: birth and death on the left and the saddle move on the right.

**Definition 4.2.** A grid movie isotopy is a sequence of grid movies such that the grid diagram in each level undergoes a grid planar isotopy.

Note that any two grid movies can only be isotopic if they have the same number of stills. However, if needed extra stills can be added by duplicating a still and inserting the extra copy right next to the original.

**4.1. Modified Planar Grid Algorithm and height functions in the stills.** In order to give a meaningful method of representing a smooth knotted surface diagram by a grid movie, we need to modify the planar grid algorithm to account for a height function. We will use a complementary coordinate system  $(v, v_1, v_2, v_3)$ . Recall that  $v$  is used to project to an  $\mathbb{R}^3$  hyperplane and  $v_1$  is the vector used to define the Morse function needed to get a smooth movie (without a fixed height function in each still). We will use  $v_2$  to define a generic height function in each still (perturb  $v_2$  if need be), which will be used to define a smooth movie with a fixed height function in each still. We will refer to this height function as the second height function, the  $v_2$  height function, or the still height function. Define the vector  $v_r$  to be a vector that is parallel to all of the vertical lines of rectangles used in the planar grid algorithm. Unfortunately, we cannot use  $v_r = v_2$  since the horizontal lines in a grid diagram would not be generic with respect to the second height function (not to mention the common occurrence of multiple intersection points along the same horizontal line). Instead we will choose  $v_r$  to be a vector in the plane containing  $v_2, v_3$  such that the angle between  $v_r$  and  $v_2$  is  $\theta < \arctan(\xi)$  for small  $\xi > 0$  (see figure 17). We will eventually choose  $\xi < \frac{1}{N}$  where  $N$  is the maximum grid index of any still of the sub-collection used to represent a knotted surface diagram. In summary, given a smooth knotted surface in  $\mathbb{R}^4$  we choose a complementary coordinate system to get a smooth movie with a fixed height function in each still and then, by choosing  $v_r$ , define the direction vertical lines are drawn in our grid diagrams.

Only two modifications need to be made to the planar grid algorithm so that it respects the height function. First, change step (1) of the planar grid algorithm to the following:

**(1h):** Draw intersection squares around each double point and draw a rectangle about each local maximum point (resp. minimum point) such that: each rectangle contains a single critical point; there are only two intersection points occurring on the bottom and right edges (resp. top and left edges) between a rectangle and the arcs of the diagram. Define these to be the critical rectangles of a grid diagram.

Secondly, when replacing the smooth arcs by grid arcs in step 4 of the planar grid algorithm, make sure that: (1) there is only corner in each non-intersection rectangle; and (2) the height ordering of the horizontal grid lines of the critical rectangles have the same height ordering as the critical points they represent.

**Definition 4.3.** *The height planar grid algorithm is the planar grid algorithm with the two modifications above.*

**4.2. Grid Movies from Smooth Movies.** We need a method of getting a grid movie from a smooth knotted surface diagram such that passing to this grid representation and back to a smooth knotted surface diagram yields a diagram isotopic to the original. In this subsection we present such a method. The details are in the constructive proof of the following theorem, which uses the dictionary found in figure 17 between basic grid diagram moves and their corresponding sequence of FESIs.

**Theorem 4.4.** *To any smooth knotted surface  $K$  a grid movie  $\mathcal{G}$  can be assigned. For any grid movie  $\mathcal{G}'$  there is a smooth knotted surface  $K'$  whose corresponding grid movie is  $\mathcal{G}'$ . Furthermore, if  $\mathcal{G} = \mathcal{G}'$  then  $K$  is smoothly isotopic to  $K'$ .*

*Proof.* We will assume familiarity with sentences and words as used in [2] for this proof. Let  $K_{[a,b]}$  be a smooth knotted surface and let  $\mathcal{K}_{[a,b]}$  be a smooth movie with a height function fixed in each still of  $K_{[a,b]}$ . By theorem 3.5.4 of [2] there is a sentence  $S$  associated to  $\mathcal{K}_{[a,b]}$ . We will assign each particular FESI or sequence of FESIs with a sequence of grid moves as indicated by the dictionary in figure 17. In order to do this we need to perform an isotopy that will change the FESIs used to represent the surface  $K_{[a,b]}$ . Specifically, act on  $\mathcal{K}_{[a,b]}$  with a level preserving isotopy of  $\mathbb{R}^3$  (that does not respect the height function) such that each occurrence of the FESIs below are sent to the sequence of FESIs indicated in figure 16.

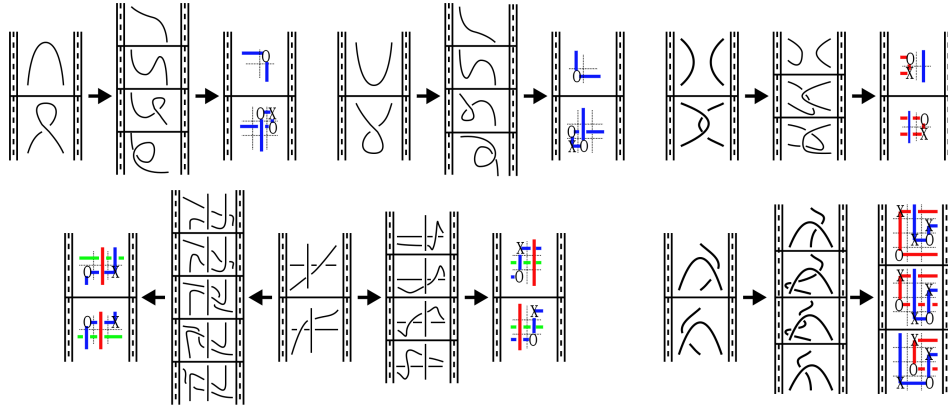


FIGURE 16. FESIs get mapped to a different sequence of FESIs, which will be identified with a basic grid move. Three other R2 FESIs exist that are similar to the top right sequence as well as three other similar transfer moves.

Let  $\mathcal{K}'_{[a,b]}$  be the knotted surface diagram resulting from this isotopy and let  $S'$  be the sentence associated to  $\mathcal{K}'_{[a,b]}$ . Except for the camel-back FESs this isotopy only induces local planar isotopies near these type of FESIs. In a camel-back FESI the movie move 23a is present. Therefore, the first movie move theorem ensures that  $\mathcal{K}_{[a,b]}$  and  $\mathcal{K}'_{[a,b]}$  represent isotopic surfaces, which means that sentences  $S$  and  $S'$  are sentences of isotopic surfaces. To get a grid movie follow the grid movie algorithm below (let  $\epsilon > 0$  be small).

**Grid Movie Algorithm:**

- (1) Apply the height planar grid algorithm to  $\mathcal{K}'_a \subset \mathbb{R}_a^2$  to get a grid diagram  $G_0$  but don't delete any quadrilaterals used, call this set of quadrilaterals  $Q_0$ . If  $\mathcal{K}'_a$  is a point then let  $G_0$  be the empty grid diagram and  $G_1$  be the grid diagram with an  $X$  and  $O$  in the only square and apply the algorithm to the still  $\mathcal{K}'_{a+\epsilon}$  to get  $G_2$ .
- (2) Slide  $Q_0$ , setting  $Q_s = Q_0$ , in the time direction until the still  $\mathcal{K}'_{a_1}$  does not satisfy the properties set forth in the height planar grid algorithm.
- (3) In the still  $\mathcal{K}'_{a_1-\epsilon}$  modify the set  $Q_0$  to the set  $Q_1$ , which will satisfy the needed conditions in  $\mathcal{K}'_{a_1}$ . If this results in more than one grid move being performed then choose an ordering for these moves and apply them individually to  $G_0$  to get the next grid diagrams in the sequence  $G_i, G_{i+1}, \dots, G_n$ .
- (4) If a critical point, of either the Morse function or the second height function, is reached then use the dictionary in figure 17 to exchange specific sequences of FESIs with the necessary basic grid move representing that type of critical point.

Specifically, for type I critical points use the corresponding grid move (birth, death, saddle). For type II and III critical points: a branch point corresponds to a Reidemeister 1 move (achieved by an R1-stabilization or R1-destabilization); a local minimum or maximum of the double point set (that is not an endpoint of  $\overline{S_2}$ ) corresponds to a Reidemeister 2 move (achieved by commutation); a triple point corresponds to a Reidemeister 3 move (achieved by a commutation). For critical points of the second height function: a cusp move is achieved by a kink stabilization or a kink destabilization; a camel back move is achieved by a transfer move; each multi-local move is achieved by a commutation.

- (5) Repeat steps 2-4 until the still  $\mathcal{K}'_b$  is reached.
- (6) Delete all sets of rectangles  $Q_i$ .

This will result in a finite sequence of grid diagrams  $\mathcal{G} = [G_0, G_1, \dots, G_m]$  for some  $m \in \mathbb{Z}$  where the brackets indicate an ordered list. Specifically,  $G_0$  corresponds to the grid diagram in stills  $\mathbb{R}_{[0, a_1-\epsilon]}^2$  generated by quadrilaterals  $Q_s$  for  $s \in [0, a_1]$  and  $G_1$  corresponds to the grid diagram in stills  $\mathbb{R}_{[a_1-\epsilon, a_2-\epsilon]}^2$  generated by quadrilaterals  $Q_s$  for any  $s \in [a_1 - \epsilon, a_2 - \epsilon)$ , etc. This sequence is a grid movie  $\mathcal{G}$  by construction and has the further property that it contains all FESIs present in sentence  $S'$  along with some additional isotopy information. Thus for any knotted surface a grid movie can be assigned to it.

Next, suppose that a grid movie  $\mathcal{G}' = [G'_1, \dots, G'_k]$  is given. Use the dictionary in figure 17 to construct a sentence  $S'$  (delete any duplicate stills to do so). Then appeal to theorem 3.5.4 in [2] to conclude that there exists a knotted surface  $K''$  with movie  $\mathcal{K}''$  whose corresponding sentence is the given one. Next, reverse the dictionary identifications to arrive at a grid movie





$\mathcal{G}''$ . In the process of creating a sentence from  $\mathcal{G}'$  any extra level preserving (and height respecting) isotopy information was deleted. Therefore,  $\mathcal{G}''$  differs from  $\mathcal{G}'$  by this isotopy information that does not effect the sentence  $S'$ . That is  $\mathcal{G}''$  differs from  $\mathcal{G}'$  by a grid movie isotopy that does not change the sentence  $S'$ . Apply a level preserving (and second height function respecting) isotopy of  $\mathbb{R}^3$  to  $\mathcal{K}''$  get a knotted surface movie  $\mathcal{K}'$  representing a surface with sentence  $S'$  that also has grid movie  $\mathcal{G}'$ . Hence, there exists a knotted surface  $K'$ , with knotted surface diagram  $\mathcal{K}'$ , whose associated grid movie is  $\mathcal{G}'$ .

For the last part of the theorem let  $K$  be a knotted surface with knotted surface diagram  $\mathcal{K}$  and associated grid movie  $\mathcal{G}$  and let  $K'$  be the knotted surface, with knotted surface diagram  $\mathcal{K}'$ , corresponding to  $\mathcal{G}$  via the process described above. We need to show that  $K$  and  $K'$  are isotopic knotted surfaces. To see this we will simply review the procedure above and make an observation. The knotted surface diagram  $\mathcal{K}$  has a sentence  $S$ , we then isotope this knotted surface diagram to an isotopic knotted surface diagram  $\mathcal{K}'$  with sentence  $S'$  and then associate a grid movie  $\mathcal{G}$  to  $\mathcal{K}'$  by the grid movie algorithm. Clearly there exists a knotted surface associated to  $\mathcal{G}$  since  $\mathcal{K}'$  gave rise to  $\mathcal{G}$ . In fact there are many different surfaces that give rise to  $\mathcal{G}$  and each of these have the same sentence  $S'$ . Therefore, any knotted surface that gives rise to  $\mathcal{G}$  must be isotopic to  $K$  since  $S$  and  $S'$  represent isotopic knotted surfaces.  $\square$

**Definition 4.5.** *The grid movie  $\mathcal{G} := [G_0, G_1, \dots, G_m]$  represents the knotted surface  $K$  and the knotted surface diagram  $\mathcal{K} = p(K)$  if  $\mathcal{G}$  could be obtained from the process in theorem 4.4*

A quick aside: it is sometimes useful to think of a grid movie as a single grid diagram together with a finite ordered list of grid moves that act sequentially on the grid diagram,

$$(2) \quad \mathcal{G} := [G_0, G_1, G_2, \dots, G_n] = [G_0, g_1 G_0, g_2 g_1 G_0, g_3 g_2 g_1 G_0, \dots, \Pi_{i=1}^n g_i G_0]$$

where  $g_k$ , for each  $k$ , is one of the following grid moves: birth, death, saddle, commutation, stabilization, destabilization, identity.

**4.3. Grid Movie Isotopies.** To get a grid movie isotopy from a smooth isotopy  $\mathcal{I}$  of a movie  $\mathcal{K}$  perform the grid movie algorithm to get a grid movie  $\mathcal{G}_0$  that represents  $\mathcal{K}_0$  but don't delete the ordered sets of rectangles  $Q_l$  for  $l \in [a, b]$ . Then run the isotopy forward until time  $t_1$  when some still  $\mathcal{K}_{b_1}$  with rectangles  $Q_{b_1}$  doesn't satisfy the properties set forth in the planar grid algorithm. Change the block of rectangles  $Q_{[b_1-c, b_1+c]}$  (for some  $c \in \mathbb{R}$ ) so that the isotopy satisfies the planar grid algorithm at  $t_1$ . This will add a grid diagram to the grid movie  $\mathcal{G}_i$  in some position  $G_k$ . Call this new grid movie  $\mathcal{G}_{i+1}$  and to make  $\mathcal{G}_{i+1}$  grid isotopic to  $\mathcal{G}_i$  insert a copy of  $G_{k-1}$  at  $G_k$  in  $\mathcal{G}_i$  (and each grid movie  $\mathcal{G}_n$  for  $n < i$ ). Repeat these steps until the isotopy is finished.

**4.4. Reidemeister, Movie, and Grid Movie Moves.** In this subsection we will discuss exactly how movie moves act on movies, develop notation that captures these changes explicitly and then apply this perspective to define grid movie moves. To accomplish this, we will begin in the common ground of Reidemeister moves acting on a diagram of a knot.

**Definition 4.6.** *A sub-diagram of a knot diagram  $E$  is a rectangular subset  $F$  of  $E$ .*

Note that a sub-diagram will usually not be a diagram of a knot. We will choose model Reidemeister moves as shown in figure 18. We denote the left and right sub-diagrams of each

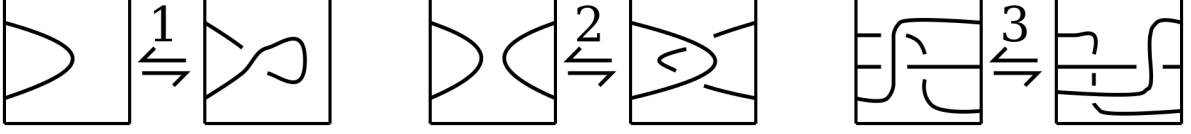


FIGURE 18. Model Reidemeister moves.

Reidemeister move by  $F_{iL}$  and  $F_{iR}$  respectively.

To be very careful, we will define a function  $R_i: E \rightarrow E'$  that operates on a knot diagram with a sub-diagram  $F_{iL}$  by defining  $R_i$  restricted to the complement of  $F_{iL}$  to be the identity and  $R_i(F_{iL}) = F_{iR}$ . Technically there are more than three moves once certain symmetries are taken into account.

Model moves are very restrictive, since there are very few diagrams that have these exact sub-diagrams. However, we are also allowed to planar isotope a diagram so that if there is a sub-diagram planar isotopic to one of the diagrams in our model moves we are still able to perform the Reidemeister move.

Let  $E$  be a diagram that has a sub-diagram that is planar isotopic (by  $I$ ) to a sub-diagram involved in a model Reidemeister move. To be very explicit, to perform a Reidemeister move on diagram  $E$  the following procedure needs to be carried out: isotope  $E$  by  $I_s$ , to get a diagram,  $I_1(E)$ , with a model Reidemeister sub-diagram; perform the Reidemeister move  $R \circ I_1(E)$ ; and then (if desired) reverse the isotopy  $\bar{I}_t \circ R \circ I_1(E)$ , where  $\bar{I}_s = I_{1-s}$ , to arrive at a diagram  $E'$ .

The movie moves are also model moves that act on a movie of a knotted surface diagram. Specifically, a movie move is a map that takes a collection of sub-diagrams  $\mathcal{D}_{iL} = \cup_{v \in [a_1, b_1]} D_v$  (where  $D_v \subset \mathbb{R}_v^2$  is a sub-diagram) to a different collection of sub-diagrams  $\mathcal{D}_{iR}$ . In notation, a movie move is a map  $\mathfrak{M}_i: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  such that  $\mathfrak{M}_i$  is the identity on the complement of  $\mathcal{D}_{iL}$  and  $\mathfrak{M}_i(\mathcal{D}_{iL}) = \mathcal{D}_{iR}$ . Again, since movie moves are model moves, they are quite restrictive and are combined with level preserving isotopies of  $\mathbb{R}^3$  (that respect the  $v_2$  height function) to make them robust. Explicitly, to perform a movie move the following procedure needs to be carried out: isotope  $\mathcal{K}_{[a,b]}$  by a level preserving isotopy  $\mathcal{I}_{[a,b]}^t$  (that respects the  $v_2$  height function), to get a movie with a model movie move; perform the movie move; and then (if desired) reverse the isotopy  $\mathcal{I}^t$  with  $\bar{\mathcal{I}}_{[a,b]}^t$  to arrive at a movie  $\mathcal{K}'_{[a,b]}$ .

We will define grid movie moves to be model moves as well. This means that a grid movie isotopy will almost always be required, since it will be rare that the exact sequence of sub-diagrams of our model move will be contained in a grid movie.

**Definition 4.7.** A sub-diagram of a grid diagram  $G$  is a rectangular collection  $D$  of grid squares from  $G$  that denotes the positions of  $O$ 's and  $X$ 's and specifies the direction of any grid arcs that intersect the boundary of  $D$ .

From now on we will use  $\mathcal{D}_{[c,d]}$  to indicate a smooth collection of sub-diagrams and  $\mathcal{D}_i$  to indicate a sequence of grid sub-diagrams. We denote smooth movie moves and grid movie moves with the same symbol  $\mathfrak{M}$ . The meaning will hopefully be clear from context. We will also denote sub-diagrams of a grid diagram by the collection of  $X$ 's and  $O$ 's and straight lines representing the knot.

**Definition 4.8.** A grid movie move  $\mathfrak{M}_i$  is a map

$$(3) \quad \mathfrak{M}_i: \mathcal{G} \rightarrow \mathcal{G}' \quad \text{defined by} \quad \begin{aligned} \mathfrak{M}_i(\mathcal{D}_{iL}) &= \mathcal{D}_{iR} \\ \mathfrak{M}_i|_{\mathcal{D}_{iL}^C} &= \text{identity}, \end{aligned}$$

where  $\mathcal{D}_{iL}^C$  is the complement of  $\mathcal{D}_{iL}$ .

We have defined grid movies, grid movie moves and grid isotopies. We have also discussed how to get grid movies and grid isotopies from smooth movies and smooth isotopies. All that remains to be done is to choose our model grid movie moves. To do this, we will simply apply the process used in theorem 4.4 to the neighborhood of a movie move. With the following definition we have finally defined all the objects needed to prove theorem 1.10.

**Definition 4.9.** The list of model grid movie moves  $\mathfrak{M}_i, i = 1 - 30$  are in the four page figure 19 and each is obtained by a particular application of the method presented in theorem 4.4 to arcs in the stills of the smooth movie moves. The last move  $\mathfrak{M}_{31}$  is discussed below.

The grid diagrammatic representation of movie move 31 can be denoted by

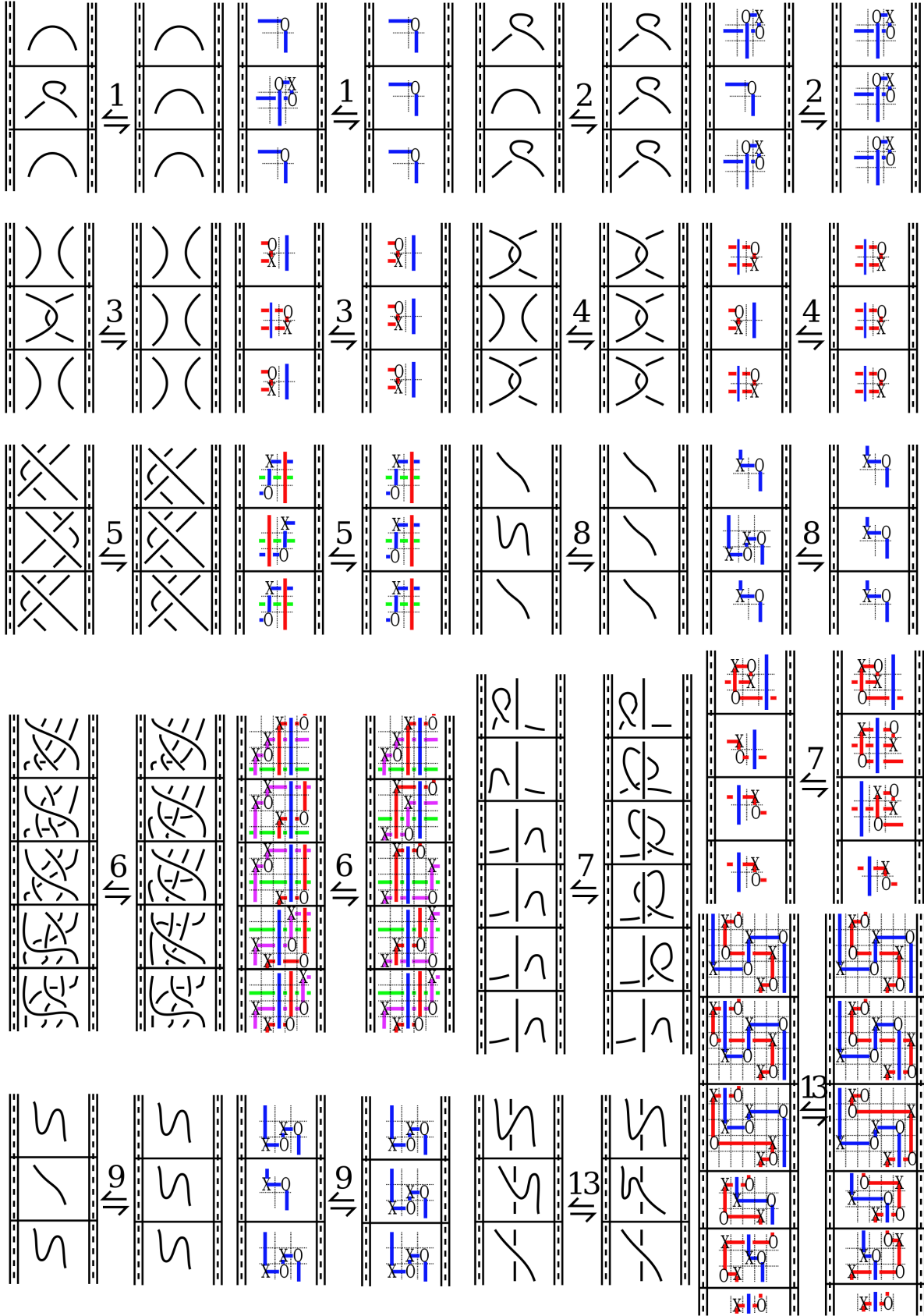
$$[Fg_1h_1E, Fg_1g_2h_1E, Fg_1g_2h_1h_2E] \leftrightarrow [Fg_1h_1E, Fg_1h_1h_2E, Fg_1g_2h_1h_2E],$$

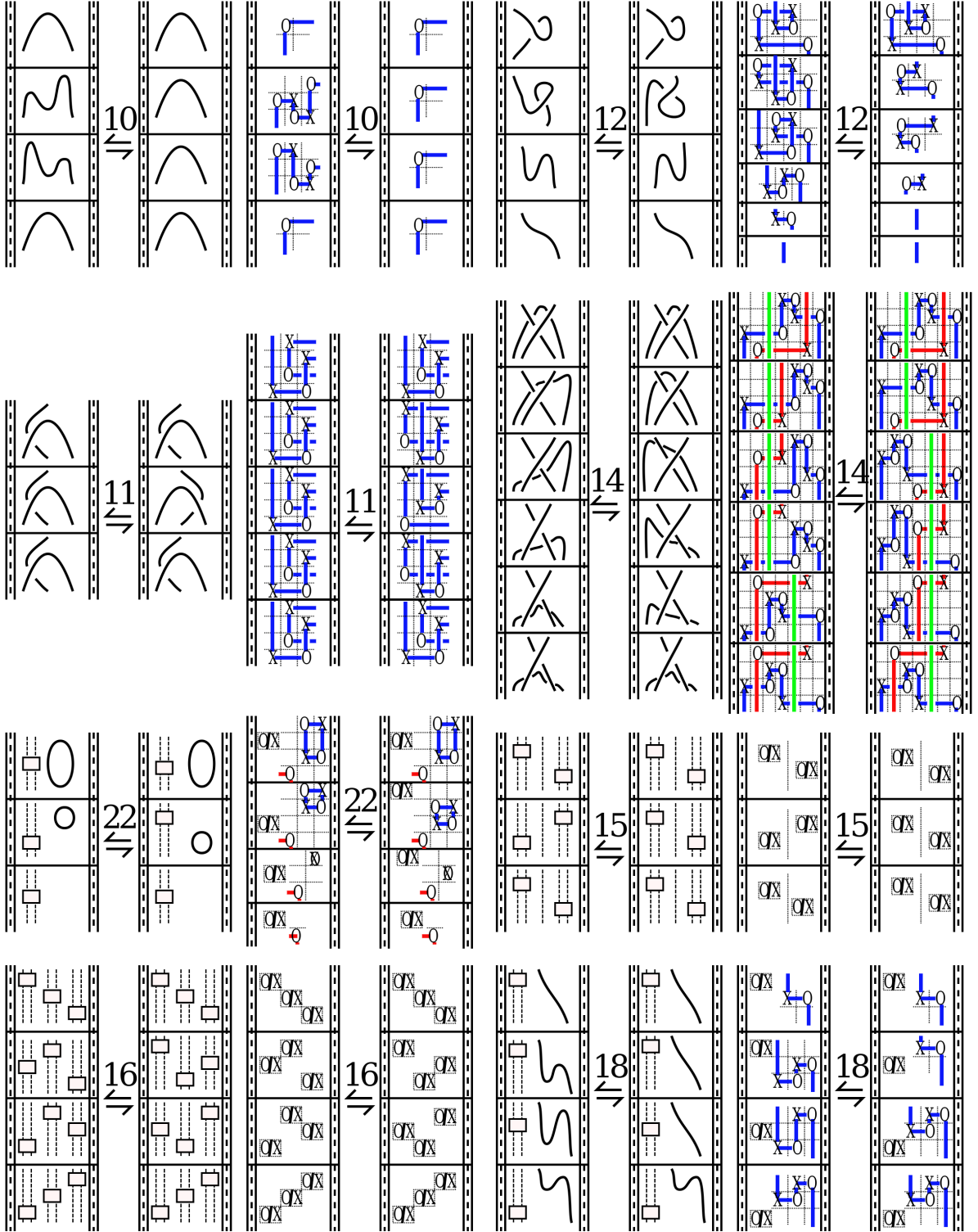
where:  $E = \prod_{i=1}^k g_i G$  is a product of grid moves  $g_i$  acting on a grid diagram  $G$ ;  $F = \prod_{j=m}^n g_j$  is a product of grid moves; and  $f_1, f_2, g_1, g_2$  are grid moves that represent FESIs.

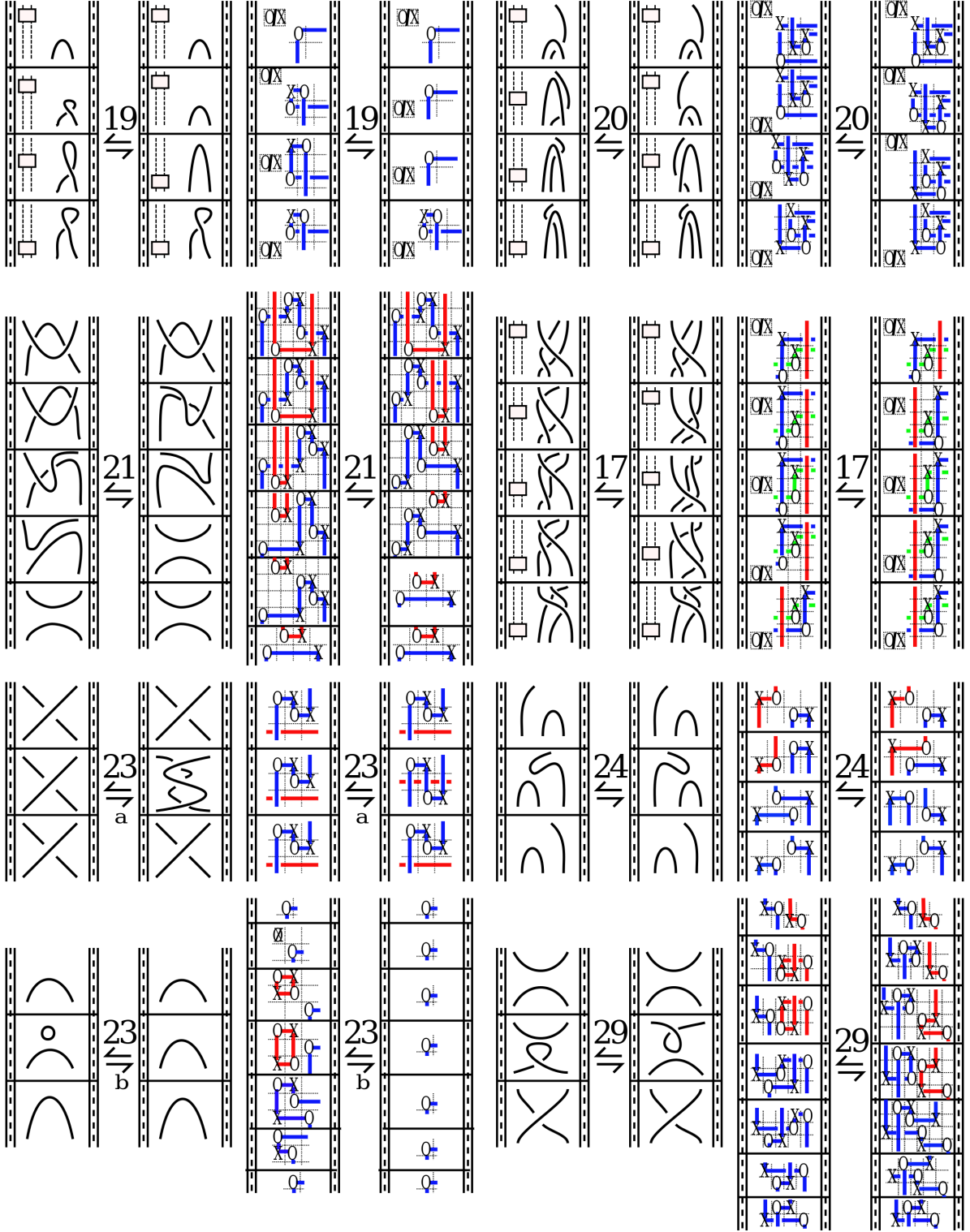
*Proof of theorem 1.10.* ( $\Rightarrow$ ) Suppose that  $\mathcal{G}_1, \mathcal{G}_2$  are grid movies that represent smooth isotopic knottings. This means that  $\mathcal{G}_i$  represents a smooth knotted surface  $K_i$  with knotted surface diagram  $\mathcal{K}_i$  for  $i = 1, 2$ . It is given that  $K_1$  and  $K_2$  are smooth isotopic knotted surfaces, which means that we can appeal to theorem 1.7 to get that  $\mathcal{K}_1$  and  $\mathcal{K}_2$  are related by a sequence of movie moves and possibly a level preserving (and second height function respecting) isotopy.

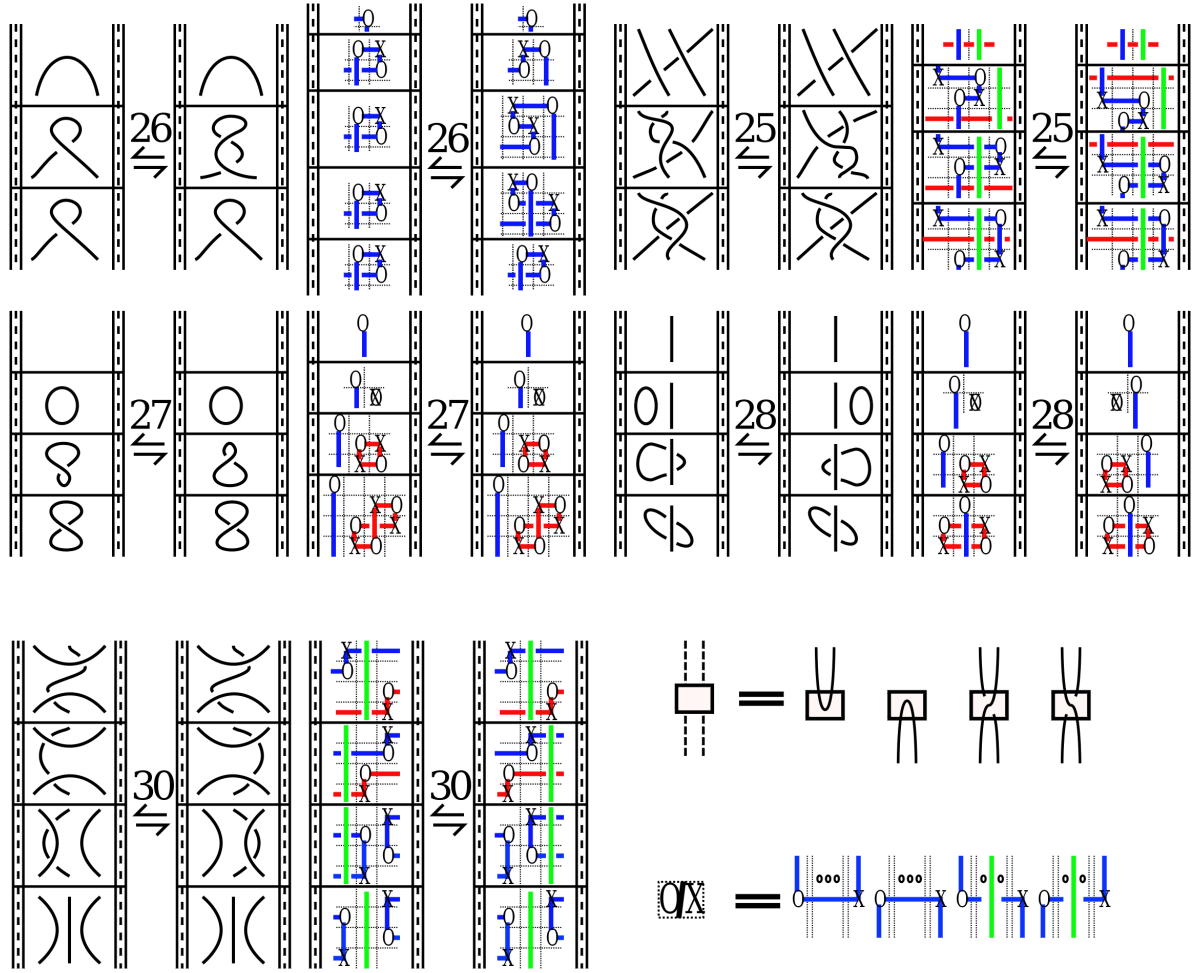
Since each model smooth movie move is represented by a model grid movie move (theorem 4.4), each smooth level preserving isotopy respecting the height function gives rise to a grid movie isotopy that respects the height function (by construction in section 4.3) and the output of the planar grid algorithm is well defined (proposition 3.10, theorem 3.11 and lemma 3.9) then  $\mathcal{G}_1$  and  $\mathcal{G}_2$  are related by a sequence of grid movie moves and possibly a grid movie isotopy respecting the height function.

( $\Leftarrow$ ) Reverse the argument given above. □







FIGURE 19. Smooth and Grid movie moves  $1, \dots, 30$ .



## 5. SURFACES WITH BOUNDARY

Our main motivation for this work is to investigate whether the combinatorial version of  $\text{HFK}^-$  is functorial over smooth marked cobordisms. The idea is to use grid movies to represent the surfaces and then compare the maps induced on the knot Floer homology chain complex by grid movies. In this paper, we have presented a generalization of Carter Rieger and Saito's movie move theorem that involves closed surfaces. The maps induced by closed surfaces, on knot Floer homology complexes, are decidedly uninteresting, since they induce a map between chain complexes with no generators. The aim of this section is to generalize the definitions and constructions used previously in this paper to account for surfaces with boundary components.

Let  $(F, \partial F)$  be a surface with boundary. Choose a vector  $v_1 \in \mathbb{R}^4$  and for each  $s \in \mathbb{R}$  let  $\mathbb{R}_s^3$  denote the hyperplane orthogonal to  $v_1$  at the point  $sv_1$ . Again, we will let  $K$  refer either to the embedding  $K: F \hookrightarrow \mathbb{R}^4$  or the image  $K(F)$  of the embedding and  $\partial K := K(\partial F)$ .

**Definition 5.1.** *An embedding of a knotted surface with boundary  $(K, \partial K, v_1, a, b)$  is an embedding  $K: F \hookrightarrow \mathbb{R}^4$  along with a choice of unit vector  $v_1 \in \mathbb{R}^4$  and  $a < b \in \mathbb{R}$  such that*

- (1)  $\partial K \subset \mathbb{R}_a^3 \cup \mathbb{R}_b^3$ ;
- (2)  $K \subset \mathbb{R}^3 \times [a, b]$ ;
- (3)  $(K - \partial K) \subset \mathbb{R}^3 \times (a, b)$ .

Note that the third condition prevents type I critical points from occurring in  $\mathbb{R}_a^3 \cup \mathbb{R}_b^3$ . We often use  $(K, \mathbb{K})$  to denote a knotted surface with boundary when the choice of  $v_1, a, b$  are not important for the discussion.

## 5.1. Boundary Knotted Surface Isotopies.

**Definition 5.2.** *Two knotted surfaces with boundary  $(K_i, \partial K_i, v_1, a, b)$ ,  $i = 1, 2$  are boundary ambiently isotopic if there is an isotopy  $H: \mathbb{R}^4 \times [0, 1] \rightarrow \mathbb{R}^4$  such that:*

- (1)  $H(x, 0) = x \quad \forall x \in \mathbb{R}^4$ ;
- (2)  $H(K_1(c), 1) = K_2(c) \quad \forall c \in F$ ;
- (3)  $H(K, s) \subset \mathbb{R}^3 \times [a, b] \quad \forall s \in [0, 1]$ ;
- (4)  $H_s|_{\mathbb{R}_a^3 \cup \mathbb{R}_b^3} = \mathbb{R}_a^3 \cup \mathbb{R}_b^3 \quad \forall s \in [0, 1]$ .

Two knotted surfaces with boundary are *fixed boundary ambiently isotopic* if the first three conditions above are satisfied and the last condition is strengthened to  $H_s|_{\mathbb{R}_a^3 \cup \mathbb{R}_b^3} = id$ .

Note that the definition of knotted surfaces with boundary and their isotopies do not allow any part of the embedded surface to intersect,  $\mathbb{R}^3 \times (-\infty, \infty) - \mathbb{R}^3 \times [a, b]$ . This prevents a “knotted” surface with boundary from “unknotting” itself by pulling components over the boundary.

There are different choices of unit vector  $v_1$  and  $a, b \in \mathbb{R}$ . Therefore in order to compare embeddings with different choices of  $v_1, a, b$  we need to describe isotopies that will make two choices of  $v_1, a, b$  coincide. In order to preserve the meaning of the vectors  $v_1$  and  $v_2$  as the vectors corresponding to two different Morse functions we will use  $w$  as our choice of unit vector

for the rest of this subsection. Suppose that  $K_i: F_i \hookrightarrow \mathbb{R}^4$  for  $i = 1, 2$  are two embeddings with choices  $w_i, a_i, b_i$ . Presently we define: a rotation isotopy  $R_{w_1 w_2}$ ; a collar isotopy  $C$ ; and a translation isotopy  $T$ , which lead to a notion of isotopy that does not depend on the choices of  $w_1, a, b$ .

**Definition 5.3.** Let  $w_1, w_2$  be two unit vectors in  $\mathbb{R}^4$ . A rotation isotopy is an isotopy  $R_{w_1 w_2}: \mathbb{R}^4 \times [0, 1] \rightarrow \mathbb{R}^4$  that rotates  $w_1$  through the plane spanned by  $w_1, w_2$  such that  $R_{w_1 w_2}(w_1, 0) = w_1$  and  $R_{w_1 w_2}(w_1, 1) = w_2$ .

A rotation isotopy applied to a set of points in  $\mathbb{R}^4$  is a rigid body rotation.

**Definition 5.4.** A collar isotopy  $C_{b, b'}$  of an embedded surface  $(K, \partial K)$  is a map  $K_{[a, b]}$  to a surface  $K_{[a, b']}$  for  $b < b'$  defined by the function  $C_{b, b'}: \mathbb{R}^3 \times \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}^3 \times \mathbb{R}$  satisfying:

$$C_{b, b'}(K_t, t, s) = \begin{cases} (K_t, t) & \forall t \in [a, b], \forall s \in [0, 1] \\ (K_b, t) & \forall t \in [b, b'], s \in [0, 1] \text{ s.t. } t \leq b + s(b' - b) \\ 0 & \text{otherwise} \end{cases}$$

The resulting smooth manifold is simply the original surface with a cylinder  $\partial K_b \times [b, b']$  glued to the boundary  $\partial K_b$ , where  $\partial K_b := K(F) \cap \mathbb{R}_b^3$ .

**Definition 5.5.** A translational isotopy  $T_{a, a'}$  of an embedded surface  $(K, \partial K)$  is a map

$$T_{a, a'}: \mathbb{R}^3 \times \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}^3 \times \mathbb{R} \text{ defined by } H(K_t, t, s) = (K_t, t + s(a' - a))$$

Translational isotopies are also rigid isotopies.

**Lemma 5.6.** Any two embedded surfaces with boundary  $(K_i, \partial K_i, w_i, a_i, b_i)$  can be isotoped to embedded surfaces with boundary contained in  $\mathbb{R}^3 \times [a_2, b_2]$  (decomposed with the vector  $w_2$ ).

*Proof.* Let  $(K_i, \partial K_i, w_i, a_i, b_i)$  be two embedded surfaces with boundary. If  $i = 2$  the second embedded surface already satisfies the requirement. So it only remains to show that we can isotope the first surface to meet the requirements. Consider the following.

$$\begin{aligned} T_{a_1, a_2} C_{b_1, b_2 - a_2 + a_1} R_{w_1 w_2}(K_1, \partial K_1, w_1, a_1, b_1) &= T_{a_1, a_2} C_{b_1, b_2 - a_2 + a_1}(K_1, \partial K_1, w_2, a_1, b_1) \\ &= T_{a_1, a_2}(K_1, \partial K_1, w_2, a_1, b_2 - a_2 + a_1) \\ &= (K_1, \partial K_1, w_2, a_2, b_2) \end{aligned}$$

We have abused notation and let  $K_1, \partial K_1$  represent the final image of each of the isotopies.  $\square$

**Definition 5.7.** Two embedded surfaces with boundary  $(K_i, \partial K_i, w_1, a_1, b_1)$  with  $i = 1, 2$  are boundary isotopic if there exists a composition of rotational, collar, translational and ambient isotopies,  $D: \mathbb{R}^4 \times [0, 1] \rightarrow \mathbb{R}^4$  satisfying:

- (1)  $D(K_1, 0) = K_1$  and  $D(\partial K_1, 0) = \partial K_1$ ;
- (2)  $D(K_1, 1) = K_2$  and  $D(\partial K_1, 1) = \partial K_2$ .

**5.2. Knotted Surface Diagrams with Boundary.** To get a knotted surface diagram of an embedded surface with boundary  $(K, \partial K, v_1, a, b)$ , choose a vector  $v \in v_1^\perp$  and project to some hyperplane  $\mathbb{R}^3$  of  $\mathbb{R}^4$ . However, the function  $p \circ K: F \rightarrow \mathbb{R}^3$  may not denote a generic surface. To fix this, perturb  $v$  slightly (while remaining in  $v_1^\perp$ ) so that  $p \circ K$  does denote a generic surface.

To get a movie of a knotted surface diagram with boundary, we simply need to check if the function of projecting onto the vector  $v_1$ ,  $p_1: \mathbb{R}^3 \rightarrow \mathbb{R}$ , is a Morse function with the desired properties. If it is then we are done. However, since we chose  $v_1$  before we chose  $v$  we may need to perturb  $v_1$  to get the required Morse function  $p_1$ . To ensure that we are still working with an embedded surface with boundary we need to perturb the boundary components of  $K$  to lie in  $\mathbb{R}_a^3$  with respect to  $v'_1$  instead  $v_1$ . Furthermore, to get a complementary coordinate system, we must perturb  $v_1$  so that  $v'_1$  remains in  $v^\perp$ . Since Morse functions are generic there are no obstructions to perturbing  $v_1$  to  $v'_1 \in v^\perp$  and since both  $v_1$  and  $v'_1$  are in  $v^\perp$  perturbing the boundary components does not effect the boundary knotted surface diagram. Finally, if a movie with a fixed height function in each still is desired, choose a generic vector  $v_2 \in \text{span}(v, v_1)^\perp$  that will be used to define the Morse height function in each still.

**Definition 5.8.** *An exceptional type II critical point is a boundary point of the double point set that is not a branch point.*

From the definition of a knotted surface with boundary, exceptional type II critical points can only occur in the projection of the boundary (of the knotted surface with boundary).

**Definition 5.9.** *A knotted surface movie with boundary is a knotted surface diagram with boundary together with a Morse function that separates all type I, II, III critical points save for the exceptional type II critical points.*

A *knotted surface movie with boundary and a height function in each still* is simply a knotted surface movie with boundary along with a second Morse function defined by projecting onto  $v_2$ .

Note that it is possible for a boundary knotted surface movie to have a critical point contained in the boundary that is not an exceptional critical point. However, this point cannot be a type I critical point.

We would like to characterize arbitrary boundary isotopies of knotted surfaces with boundary. This requires us to enlarge the movie moves by adding boundary movie moves that will account for the changes in the boundary induced by boundary isotopies. Fortunately, the needed boundary moves are a subset of the FESIs.

**Definition 5.10.** *The boundary movie moves are comprised of FESIs that do not represent type I critical points (births, saddles and deaths).*

Let  $(K, \partial K)$  be a knotted surface with boundary that is represented by the knotted surface movie with boundary  $(\mathcal{K}, \partial \mathcal{K})$ . Given the definition of a knotted surface with boundary, type I critical points cannot occur in  $\partial \mathcal{K}$ . However, every other type of critical point can occur in the boundary, even though this is not the generic situation. The types of boundary isotopies

that induce a non-exceptional critical point in  $\partial\mathcal{K}$  during the isotopy are cataloged by the FESIs that do not represent type I critical points (births, deaths, and saddles). Specifically, an arbitrary boundary isotopy can induce the following singularities in a boundary component of  $\partial\mathcal{K}$ :

- an addition or deletion of a branch point of  $(\mathcal{K}, \partial\mathcal{K})$  that induces a Reidemeister one move in  $\partial\mathcal{K}$ ;
- a birth or death of a double point line in the movie that induces a Reidemeister two move in  $\partial\mathcal{K}$ ;
- an addition or deletion of a triple point that induces a Reidemeister three move in  $\partial\mathcal{K}$ ;
- a singularity of the second height function that induces a multi-local move in  $\partial\mathcal{K}$ .

**Theorem 5.11.** *Two knotted surface movies with boundary, with a height function fixed in each still, represent boundary isotopic knottings if and only if they are related by a finite sequence of movie moves and/or boundary movie moves.*

*Proof.* Let  $(K_i, \partial K_i, v_i, a_i, b_i)$  for  $i = 1, 2$  be two boundary isotopic knotted surfaces with boundary. Using lemma 5.6 we can isotope these surfaces to  $(K_i, \partial K_i, v_1, a, b)$  for some  $v_1, a, b$ . We need to show that  $(\mathcal{K}_i, \partial\mathcal{K}_i, v_1, v_2, a, b)$  for  $i = 1, 2$ , the movies (with fixed height functions in each still) representing these knotted surfaces with boundary, are related by a finite sequence of movie moves and boundary movie moves iff  $(K_1, \partial K_1)$  and  $(K_2, \partial K_2)$  are boundary ambiently isotopic. Further suppose that  $p: \mathbb{R}^4 \rightarrow \mathbb{R}^3$  is the generic projection used in the movie description.

First assume that  $(K_i, \partial K_i)$  for  $i = 1, 2$  are related by a boundary isotopy  $I: \mathbb{R}^4 \times [0, 1] \rightarrow \mathbb{R}^4$  and assign sentences  $S_1, S_2$  to the movies  $(\mathcal{K}_1, \partial\mathcal{K}_1), (\mathcal{K}_2, \partial\mathcal{K}_2)$ . The boundary isotopy can either induce singularities, of the projection, in the interior or the boundary of  $p \circ I(\mathcal{K}, s)$ . By appealing to the enumeration of singularities found in the proof of the smooth movie theorem [2] we know that if a singularity occurs in the interior of  $p \circ I(\mathcal{K}_1, s)$ , for some  $s \in [0, 1]$ , then the isotopy will induce a change in the sentence  $S_1$  that corresponds to a movie move. If the singularity occurs on the boundary, following the enumeration of possible boundary singularities above, this will induce a change in the sentence corresponding to one of the boundary movie moves. Hence, there exists a sequence of movie moves and boundary movie moves that relate  $S_1$  to  $S_2$  (and therefore  $(\mathcal{K}_1, \partial\mathcal{K}_1)$  to  $(\mathcal{K}_2, \partial\mathcal{K}_2)$ ).

The fact that for each movie move and boundary movie move there exists an isotopy of the movie that induces the required move allows us to reverse the proof.  $\square$

If desired one could work with a definition of boundary isotopy that required the isotopy of the boundary be a planar isotopy.

**Definition 5.12.** *Planar boundary movie moves consist of the multi-local move FESIs: all FESIs that do not represent type I, II or III critical points.*

**Corollary 5.13.** *Two knotted surfaces with boundary  $(K_i, \partial K_i)$  for  $i = 1, 2$  with movies (with a height function fixed in each still)  $(\mathcal{K}_i, \partial\mathcal{K}_i)$  are related by a boundary isotopy that induces a planar isotopy in  $\partial\mathcal{K}_i$  if and only if they are related by a finite sequence of movie moves and planar boundary movie moves.*

*Proof.* The proof follows immediately from the theorem by noting that since the boundary isotopy must induce a planar isotopy in the boundary of the movie, no critical points of type I, II or III can occur in the boundary. This only leaves the possibility of critical points of the second height function occurring in the boundary, which induce the planar boundary movie moves defined above.  $\square$

Finally, one could consider boundary isotopies that leave the boundary fixed.

**Corollary 5.14.** *Two boundary knotted surface movies (with a height function fixed in each still) are related by a boundary isotopy that fixes the boundary if and only if they are related by a finite sequence of movie moves.*

*Proof.* This follows from the main theorem by observing that if the isotopy fixes the boundary then no singularities, except for exceptional type II singularities will occur in the projection of the boundary. This means that all singularities induced by the isotopy must occur in the interior of the movie. Each of these singularities give rise to a movie move.  $\square$

**Definition 5.15.** *A grid movie with boundary,  $(\mathcal{G}, \partial\mathcal{G})$ , represents the knotted surface with boundary  $(K, \partial K)$  and the knotted surface diagram with boundary  $(\mathcal{K}, \partial\mathcal{K})$  if  $(\mathcal{G}, \partial\mathcal{G})$  could be obtained from  $(\mathcal{K}, \partial\mathcal{K})$  from the process in theorem 4.4.*

**Definition 5.16.** *Boundary grid movie moves consist of sequences of grid diagrams found in the dictionary (figure 17) that do not correspond to a single smooth still and that do not represent type I critical points.*

**Theorem 5.17.** *Two grid movies with boundary represent isotopic knotted surfaces with boundary if and only if they are related by a sequence of grid movie moves and/or boundary grid movie moves.*

*Proof.* Let  $(\mathcal{G}_1, \partial\mathcal{G}_1)$  and  $(\mathcal{G}_2, \partial\mathcal{G}_2)$  represent isotopic knottings with boundary  $(K_i, \partial K_i)$  with movies  $(\mathcal{K}_i, \partial\mathcal{K}_i, v_1, v_2, a, b)$  for  $i = 1, 2$  (that have a fixed height function in each still). Since  $(K_1, \partial K_1)$  is boundary isotopic to  $(K_2, \partial K_2)$  theorem 5.11 implies that there exists a finite sequence of movie moves and/or boundary movie moves relating  $(\mathcal{K}_1, \partial\mathcal{K}_1, v_1, v_2, a, b)$  and  $(\mathcal{K}_2, \partial\mathcal{K}_2, v_1, v_2, a, b)$ . For each one of these movie moves and boundary movie moves there is a corresponding grid movie move or boundary grid movie move gotten from the dictionary in figure 17. Therefore,  $(\mathcal{G}_1, \partial\mathcal{G}_1)$  and  $(\mathcal{G}_2, \partial\mathcal{G}_2)$  are related by a finite sequence of grid movie moves and/or boundary grid movie moves. Reverse the argument to get the other direction.  $\square$

**Definition 5.18.** *Planar boundary grid movie moves consist of sequences of grid diagrams found in the dictionary (figure 17) that do not correspond to a single smooth still and that do not represent type I, II or III critical points.*

Note that transfer moves represent camel-back moves, and therefore are a planar boundary grid move, even though they contain a birth and death of a double point arc.

**Corollary 5.19.** *Two grid movies with boundary represent two knottings with boundary  $(K_i, \partial K_i)$ ,  $i = 1, 2$  with movies  $(\mathcal{K}_i, \partial\mathcal{K}_i, v_1, v_2, a, b)$  that are related by a boundary isotopy*

that induces a planar isotopy in  $\partial K_i$ , if and only if they are related by a sequence of grid movie moves and/or boundary grid movie moves.

*Proof.* This follows from theorem 5.17 and corollary 5.13.  $\square$

**Corollary 5.20.** *Two grid movies with fixed boundary represent isotopic knotted surfaces with fixed boundary if and only if they are related by a sequence of movie moves.*

*Proof.* This follows from theorem 5.17 and corollary 5.14.  $\square$

## REFERENCES

1. R. H. Bing, *The geometric topology of 3-manifolds*, American Mathematical Society Colloquium Publications, vol. 40, American Mathematical Society, Providence, RI, 1983. MR 728227 (85j:57001)
2. J. Scott Carter, Joachim H. Rieger, and Masahico Saito, *A combinatorial description of knotted surfaces and their isotopies*, Adv. Math. **127** (1997), no. 1, 1–51. MR 1445361 (98c:57023)
3. J. Scott Carter and Masahico Saito, *Reidemeister moves for surface isotopies and their interpretation as moves to movies*, J. Knot Theory Ramifications **2** (1993), no. 3, 251–284. MR 1238875 (94i:57007)
4. Peter R. Cromwell, *Embedding knots and links in an open book i: Basic properties*, Topology Appl. **64** (1995), 37–58.
5. I. A. Dynnikov, *Arc-presentations of links: monotonic simplification*, Fund. Math. **190** (2006), 29–76. MR 2232855 (2007e:57006)
6. R. H. Fox, *A quick trip through knot theory*, Topology of 3-manifolds and related topics (Proc. The Univ. of Georgia Institute, 1961), Prentice-Hall, Englewood Cliffs, N.J., 1962, pp. 120–167. MR 0140099 (25 #3522)
7. Matthew Graham, *Marked Movies*, pre-print (2013).
8. Andras Juhász, *Cobordisms of sutured manifolds*, eprint <http://arxiv.org/abs/0910.4382v2> (2009).
9. Ciprian Manolescu, Peter Ozsváth, and Sucharit Sarkar, *A combinatorial description of knot Floer homology*, Ann. of Math. (2) **169** (2009), no. 2, 633–660. MR 2480614 (2009k:57047)
10. Ciprian Manolescu, Peter Ozsváth, Zoltán Szabó, and Dylan Thurston, *On combinatorial link Floer homology*, Geometry and Topology, 11:2339–2412, 2007 (2006).
11. Ciprian Manolescu, Peter Ozsváth, and Dylan Thurston, *Grid diagrams and Heegaard Floer invariants*, eprint <http://arxiv.org/abs/0910.0078v1> (2009).
12. Peter Ozsváth, Andras Juhász, and Dylan Thurston, *in preparation*.
13. Peter Ozsváth and Zoltán Szabó, *Holomorphic disks and knot invariants*, Adv. Math. **186** (2004), no. 1, 58–116. MR 2065507 (2005e:57044)
14. ———, *Holomorphic disks and topological invariants for closed three-manifolds*, Ann. of Math. (2) **159** (2004), no. 3, 1027–1158. MR 2113019 (2006b:57016)
15. Jacob Rasmussen, *Floer homology and knot complements*, Ph.D. thesis, Harvard University, 2003.
16. Dennis Roseman, *Reidemeister-type moves for surfaces in four-dimensional space*, Knot theory (Warsaw, 1995), Banach Center Publ., vol. 42, Polish Acad. Sci., Warsaw, 1998, pp. 347–380. MR 1634466 (99f:57029)
17. Sucharit Sarkar, *Grid diagrams and the Ozsváth-Szabó tau-invariant*, eprint <http://arxiv.org/abs/1011.5265v1> (2010).
18. ———, *Moving basepoints and the induced automorphisms of link Floer homology*, eprint <http://arxiv.org/abs/1109.2168v1> (2011).

*E-mail address:* matthew.graham.2@bc.edu

DEPARTMENT OF MATHEMATICS, BOSTON COLLEGE, CHESTNUT HILL, MA 02467-3806